

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen, Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2001, 01:34:49 ; Search time 14302.9 Seconds
(without alignments)
1528.577 Million cell updates/sec

Title: US-09-294-298-1
Perfect score: 4272
Sequence: 1 atgtcctatgcccccttcag.....tgggggtgctatcccccatcct(4272)

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues
Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- GenEmbl:*
 - 1: gb_bal:*
 - 2: gb_ba2:*
 - 3: gb_om:*
 - 4: gb_ov:*
 - 5: gb_ph:*
 - 6: gb_pl1:*
 - 7: gb_pl2:*
 - 8: gb_pr1:*
 - 9: gb_pr2:*
 - 10: gb_pr3:*
 - 11: em_fun:*
 - 12: em_hum1:*
 - 13: em_hum2:*
 - 14: em_in:*
 - 15: em_om:*
 - 16: em_or:*
 - 17: em_ov:*
 - 18: em_pat:*
 - 19: em_ph:*
 - 20: em_pl:*
 - 21: em_ro:*
 - 22: em_sts:*
 - 23: em_sy:*
 - 24: em_un:*
 - 25: em_vi:*
 - 26: gb_htg1:*
 - 27: gb_htg2:*
 - 28: gb_in1:*
 - 29: gb_in2:*
 - 30: em_bal:*
 - 31: em_ba2:*
 - 32: em_hum3:*
 - 33: em_hum4:*
 - 34: gb_pr4:*
 - 35: gb_htg3:*
 - 36: gb_htg4:*
 - 37: gb_htg5:*
 - 38: gb_htg6:*
 - 39: gb_htg7:*
 - 40: em_htg1:*
 - 41: em_htg2:*
 - 42: em_htg3:*
 - 43: em_hum5:*

- 44: gb_pl3:*
- 45: gb_pr5:*
- 46: gb_htg8:*
- 47: gb_htg9:*
- 48: gb_htg10:*
- 49: gb_htg11:*
- 50: gb_htg12:*
- 51: gb_htg13:*
- 52: gb_htg14:*
- 53: gb_in3:*
- 54: gb_htg15:*
- 55: gb_htg16:*
- 56: gb_htg17:*
- 57: em_htg4:*
- 58: em_htg5:*
- 59: em_htg6:*
- 60: em_htg7:*
- 61: em_hum6:*
- 62: gb_htg18:*
- 63: gb_htg19:*
- 64: gb_ba3:*
- 65: em_htg8:*
- 66: em_htg9:*
- 67: em_htg10:*
- 68: gb_pr6:*
- 69: gb_pr7:*
- 70: gb_htg20:*
- 71: gb_htg21:*
- 72: gb_htg22:*
- 73: gb_htg23:*
- 74: gb_ro:*
- 75: gb_sts1:*
- 76: gb_sts2:*
- 77: gb_sy:*
- 78: gb_un:*
- 79: gb_vil:*
- 80: gb_vil2:*
- 81: gb_pat1:*
- 82: gb_pat2:*
- 83: em_htg0:*
- 84: gb_htg24:*
- 85: gb_pr8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4167.4	97.6	4539	74	AF058789	AF058789 Rattus no
2	4018.2	94.1	4140	74	AF058790	AF058790 Rattus no
3	3940.4	92.2	4063	74	AF048976	AF048976 Rattus no
4	3919	91.7	4265	74	AF050183	AF050183 Rattus no
5	3602.8	84.3	4801	74	AB016962	AB016962 Rattus no
6	920.8	21.6	94770	69	HSDJ570F3	AL050332 Human DNA
7	920.8	21.6	150956	70	AL161903	AL161903 Homo sapi
8	920.8	21.6	169323	70	AL161906	AL161906 Homo sapi
9	627.8	14.7	4368	10	AF047711	AF047711 Homo sapi
10	530	12.4	829	74	AF053938	AF053938 Rattus no
11	336.4	7.9	4287	34	AK024488	AK024488 Homo sapi
12	241.8	5.7	56366	39	AC019800	AC019800 Drosophila
13	241.8	5.7	191504	36	AC012162	AC012162 Drosophila
14	241.8	5.7	202741	36	AC012161	AC012161 Drosophila
15	241.8	5.7	300994	28	AE003506	AE003506 Drosophila
16	236.8	5.5	3984	28	AB011280	AB011280 Caenorhab
17	223.6	5.2	65961	38	AC016523	AC016523 Drosophila
18	205.2	4.8	2385	34	AK022662	AK022662 Homo sapi
19	196	4.6	172027	72	AL365274	AL365274 Homo sapi
20	194.4	4.6	211	74	AF055883	AF055883 Rattus no
21	187.6	4.4	169323	70	AL161906	AL161906 Homo sapi

CHEU
1997

Db	721	 AACAGCCGCGGTAGATAACGTGCTGAAACTATGGATCATAGAAGCTCGAGAGTGCCCC	780
Qy	781	 cccaagaagcgatattactgcgagttatgcctggacgacatgctctatgcacggaccact	840
Db	781	 CCCAAGAAACGATATTACTCGGAGTTATGCCCTGGACACATGCTCTATGCACGGACCACT	840
Qy	841	 tccaaagccccgctcagccctcaggagacactgtcttttggggcgagcacttcgagtttaac	900
Db	841	 TCCAAGCCCCGCTCAGCCTCAGGAGACACTGTCTTTTGGGGCGAGCACTTCGAGTTTAAC	900
Qy	901	 aacctgcctgctgtccgggcgctgcgctgcatctgtacogtgactcggacaaaaagcgg	960
Db	901	 AACCTGCTGTGTCCGGGCGCTGCGGCTGCATCTGTACCGTGACTCGGACAAAAGCGG	960
Qy	961	 aagaaggacaaggcaggtacgttgccctggtgactgttcagtgccaccctcggtcggg	1020
Db	961	 AAGAAGGACAAGGCAGGCTACGTTGGCCTGGTGACTGTTCCAGTGGCCACCCCTGGCTGGG	1020
Qy	1021	 cgccacttcacagagcagtggttaccctcgtagccctgccaacagggaagtgggggcctcctggg	1080
Db	1021	 CGCCACTTCACAGAGCAGTGTACCCCGTGACCCTGCCAACAGGAAGTGGGGCTCTGGG	1080
Qy	1081	 ggtaggggctcgggggagggggggggtcagggggcggtcaggggggcaaaagggaagga	1140
Db	1081	 GGTATGGGCTCGGGGGGAGGAGGGGGGTACGGGGCGGCTCAGGGGGCAAGGGAAAGGA	1140
Qy	1141	 ggctgctcgtgtgcggctgaagggcccggttacccagacaatgagtatcctgccccatggag	1200
Db	1141	 GGCTGTCTGCTGTGCGGCTGAAGGCCCGGTTACCAGACAATGAGTATCCTGCCCATGGAG	1200
Qy	1201	 ctataaaggagtttgcagaaatatgtaccaaaccaactaccgcatgctgtgtgccgtgctg	1260
Db	1201	 CTATAAAGGAGTTTGCAGAAATATGTGACCAACCACCTACCGCATGCTGTGTGCCGTGCTG	1260
Qy	1261	 gagcccgccctcaatgtcaagggcaaggaggaggtcgctagtgcactggttcacatcctg	1320
Db	1261	 GAGCCCGCCCTCAATGTCAAGGGCAAGGAGGAGGTGCTAGTGCACCTGGTTACATCCTG	1320
Qy	1321	 caaagcacaggccaaggccaaggaacttcctttcagacatggcccatgtcagaggtagaccgg	1380
Db	1321	 CAAAGCACAGGCAAGGCCAAGGACITTCCTTTCAGACATGGCCATGTGAGAGGTAGACCGG	1380
Qy	1381	 ttcatggagggggaacacctcatatttcgcgcagaaacacgctcgccactaaagccatagaa	1440
Db	1381	 TTCATGGAGCGGGAACACCTCATATTCCGGGAGAACACGCTCGCCCACTAAAGCCATAGAA	1440
Qy	1441	 gagtatatgagactgattggccagaaaaatacctcaaggatgccattggggagttcatccgg	1500
Db	1441	 GAGTATATGAGACTGATTGGCCAGAAATACCTCAAGGATGCCATTGGGGAGTTTCATCCGG	1500
Qy	1501	 gctctgtatgaatctgaggagaactgtgaagttagacccccatcaagtgcacagcgtccagt	1560
Db	1501	 GCTCTGTATGAATCTGAGGAGAACTGTGAAGTAGACCCCACTCAAGTGCACACGGTCCAGT	1560
Qy	1561	 ctggcagagcaccaggccaacctgcggatgtgctgtgagttggccctgtgcaaggtggtc	1620
Db	1561	 CTGGCAGAGCACCAAGGCCAACCCTGCGGATGTGCTGTGAGTTGGCCCTGTGCAAGGTGGTC	1620
Qy	1621	 aactccattgcgtgttccccgagggagcgtgaaggagggtgtttgcatcatcgtggcggtcgc	1680
Db	1621	 AACTCCCATTCGTGTTCCCGAGGGAGCTGAAGGAGGTGTTTGCATCATGSGCGCTGCCG	1680
Qy	1681	 tgtgcagagcggggccggaggacattgctgcagggctgatcagcgcctcgctcttcctg	1740
Db	1681	 TGTGCAGAGCGGGGCGGGGAGGACATTGCTGACAGGCTGATCAGCGCCTCGCTCTTCCTG	1740
Qy	1741	 cgcttcctctgcgcggccatcatgtcgccccagttctgtttggactgatgcagagtagtaccac	1800
Db	1741	 CGCTTCCTCTGCCCCGGCCATCATGTGCCCCAGTCTGTTTGGACTGATGCAGGAGTACCCA	1800
Qy	1801	 gatgagcagacctcaggaacccctcacccctcatcgccaagggttatccagaaacctggccaac	1860

Db	1801	GATGAGCAGACCTCACGAAACCTCACCTCATCGCCAAGGTTATCCAGAAACCTGGCCAAC	1860
Qy	1861	ttttcaaagtttacctcaaaaggaggaacttcctgggcttcatagaacgagtttctggaactg	1920
Db	1861	TTTTCCAAGTTTACCTCAAAGGAGGACTTCCTGGGCTTCATGAACGAGTTTCTGGAGCTG	1920
Qy	1921	gagtggggttctatgcagcaattcttgtatgagatatccaaacctggacacactgaccaaac	1980
Db	1921	GAGTGGGGTTCTATGCAGCAATTCTTGATGAGATATCCAACCTGGACACACTGACCAAC	1980
Qy	1981	agcagcagtttgagggtacatagacttggccgcgagctctccacacttcacgcctg	2040
Db	1981	AGCAGCAGTTTGAGGGCTACATAGACTTGGGCCCGGAGCTCTCCACACTTCACGCCCTG	2040
Qy	2041	ctctgggaggtgctgccccagctcagcaaggaagccctcctgaagctgggcccgcctgccc	2100
Db	2041	CTCTGGGAGTGTCTGCCCCAGCTCAGCAAGGAAGCCCTCCTGAAGTGGGCCCGCTGCC	2100
Qy	2101	cggctcctcagcgacatcagcacagccctgaggaaaccttaacatccaaaggcagccgagc	2160
Db	2101	CGGCTCCTCAGCGACATCAGCACAGCCCTGAGGAACCTTAACATCCAAAGGCAGCCGAGC	2160
Qy	2161	cgccagagcgagcgctcggtctcagcccatggtgctgcgcgggcgcgcagccgagatg	2220
Db	2161	CGCCAGAGCGAGCGCTCGGTCTCAGCCCATGGTGTCTGCGGGGCCGTGAGCCGAGATG	2220
Qy	2221	cagggtacatgatcggggacctcaacagctcccatcgaccttcagtcttcattgctcga	2280
Db	2221	CAGGGCTACATGATCGGGGACCTCAACAGCTCCATCGACCTTCAGTCTCTCATGGCTCGA	2280
Qy	2281	ggcctcaacagctctatggacatggctcgctgcctcccccccccaacaaaggagaaacccccg	2340
Db	2281	GGCCTCAACAGCTCTATGGACATGGCTCGCTCGCTCCCTCCCCAACCAAGGAGAAACCCCG	2340
Qy	2341	ccgccccctccggtgggggtaaagacctgttctatgtgagccggcccacactggccccgg	2400
Db	2341	CGCCCCCTCCCGGTGGGGGTAAGACCTGTTCTATGTAGCCGGCCACCACTGGCCCCGG	2400
Qy	2401	tcctcccagcatactgcacgagcagctcggacatcacagagccggagcagaagatgctg	2460
Db	2401	TCCTCCCCAGCATATGACGAGCAGCTCGGACATCACAGAGCCGGAGCAGAGATGCTG	2460
Qy	2461	agtgtcaacaagagtgtgtccatgctggacctgcagggcgacggcctgggggcccgcctt	2520
Db	2461	AGTGTCAACAAGAGTGTGTCCATGCTGGACCTGCAAGGGCGACGGGCTTGGGGGGCCGCTT	2520
Qy	2521	aacagcagtagtgtttccaaacctggcagctgttggggacctgttgcaactcaagccaggct	2580
Db	2521	AACAGCAGTAGTGTTTCCAACCTGGCAGCTGTTGGGACCTGTTGCACCTCAAGCCAGGCT	2580
Qy	2581	tcactgacagcagccttgggggtgcgcctgcacctgcgggcgcctctcccaaggagat	2640
Db	2581	TCACTGACAGCAGCCTTGGGTTGCGGCTGCACCTGCGGGCGCCTCTCCCAAGGAGT	2640
Qy	2641	ggctcttccatcacagcagccggcatgcgcctcagccagatgggtgcactacggatgggt	2700
Db	2641	GGCTCTTCCATCACAGCAGCCGGCATGCGCCTCAGCCAGATGGGTGTCACTACGGATGGT	2700
Qy	2701	gtccccgccagcaactgcgcataccctcttcttccagaacctctcttccatatggct	2760
Db	2701	GTCGCCGCCCAGCAACTGGCGATCCCTCTTTCTCTCCAGAACCTCTCTTCCATATGGCT	2760
Qy	2761	gccgatggaccagggccccccagcaggccatggagggagcagtgccatgggtccacctcc	2820
Db	2761	GCCGATGGACCAGGGCCCCCAGCAGGCCCATGGAGGGAGCAGTGGCCCATGGTCCACCTTCC	2820
Qy	2821	tcccatcaccaccaccaccatcaccatcacccgagggggagaaacccccaggggagact	2880
Db	2821	TCCCATCACCAACCAACCAACCATCACCATCACCCGAGGGGGAGAACCCCCAGGGGACACT	2880
Qy	2881	tttgccccgttccatggctatagcaagagcgaggacctctctctacagggggtccctaagccc	2940
Db	2881	TTTGCCCCGTTCATGGCTATAGCAAGAGCGGAGGACCTCTCTACAGGGGTCCCTAAGCCC	2940

Db 1918 GTGCTGCCCGAGCTCAGCAAGGAAGCCCTCCTGAAGCTGGGCCCGCTGCCCGCTCCTC 1977
QY 2110 agcgacatcacagaccctgaggaaccctaacatccaaaggcagccgagccgccagagc 2169
Db 1978 AGCGACATCAGCACAGCCCTGAGGAACCCCTAACATCCAAAGGACGCCGAGCCGCCAGAGC 2037
QY 2170 gagcgctcggtctcagcccatggtgctgcgcgggcccgtcagccgagatgcagggctac 2229
Db 2038 GAGCGCGCTCGGTCTCAGCCCCATGGTGTGCGCGGCCGTACGCCGAAATGCAGGGCTAC 2097
QY 2230 atgatggggacctcaacagctccatcgaccttcagctccttcagtccttcagtcgagccctcaac 2289
Db 2098 ATGATGCGGGACCTCAACAGCTCCATCGACCTTCAGTCCCTTCATGGCTCGAGGCCTCAAC 2157
QY 2290 agctctatggacatggctcgctcgctccctccctcccaaccaaggagaaaccccccgccccct 2349
Db 2158 AGCTCTATGGACATGGCTCGCCTCCCTCCCAACCAAGGAGAAACCCCGCGCCCCCT 2217
QY 2350 cccggtggggtaagacctgttctatgtgagccggccaccactggcccggtcctcccca 2409
Db 2218 CCCGCTGCGGGTAAAGACCTGTCTATGTGAGCCGGCCACCACTGGCCCCGTCTCCCCA 2277
QY 2410 gcatactgcacgagcagctcgacatcacagagccggagcagaagatgctgagtgtaac 2469
Db 2278 GCATACTGCACGAGCAGCTCGACATCACAGAGCCGGAGCAGAAGATGCTGAGTGTCAAC 2337
QY 2470 aagagtgttccatgctggacctgcagggcgacggcgctggggcgcccttaacagcagt 2529
Db 2338 AAGAGTGTGCCATGCTGGACTGCAGGGCGACGGGCCCTGGGGGCCCGCTTAACAGCAGT 2397
QY 2530 agtgtttccaaacctggcagctggtgggacctgttgacccaagccagaccttcactgaca 2589
Db 2398 AGTGTTTCCAACTGGCAGCTGTGGGACCTGTTGCGACCTCAAGCCAGGCTTCACTGACA 2457
QY 2590 gagccttggggttgccgtgcacctgcagctgccggcgccctcccaaggagtggtcttcc 2649
Db 2458 GCAGCCTTGGGTTTGGGCCCTGCACCTGCCGGCGCCTTCCCAAGGGAGTGGCTCTTCC 2517
QY 2650 atcacagcagccggcatgcgctcagccagatgggtgtcactacggaatgggtcccccgcc 2709
Db 2518 ATCACAGAGCCCGCATGCGCCTCAGCCAGATGGGTGTCACTACGGATGGTGTCCCCGCC 2577
QY 2710 cagcaactgcgcacctcttcttccagaaacctctcttccatatggtgctgcagatgga 2769
Db 2578 CAGCAACTGCGCATCCCTCTTCTTCCAGAACCTCTCTTCCATATGGCTGCCGATGGA 2637
QY 2770 ccagggtccccagcaggccatggaggagcagtggtccatgggtccaccttcccatcac 2829
Db 2638 CCAGGGCCCCCAGCAGGCCATGGAGGGAGCAGTGGCCATGGTCCACCTTCTTCCCATCAC 2697
QY 2830 caccaccaccaccatcacatcacagcagggagacccctctctacaggggtccctaaagccccctgcgcc 2889
Db 2698 CACCACCACCACCATCACCATCACCGAGGGGAGAACCCCCCAGGGGACACTTTTGGCCCCG 2757
QY 2890 ttccatggctatagcaagagcagggacctctctacaggggtccctaaagccccctgcgcc 2949
Db 2758 TTCCATGGCTATAGCAAGAGCGAGGACCTCTCTACAGGGGTCCCTAAGCCCCCTGCGGCC 2817
QY 2950 tccatccttcacagccacagctacagtgatgagtttggacctctggtactgatatttacc 3009
Db 2818 TCCATCTTCAAGCCACAGCTACAGTGATGAGTTTGGACCCCTCTGGTACTGATTTTACC 2877
QY 3010 cgtcggcagctctcacttcaggagaaacctcacagacatgctctcccccgccccagatcac 3069
Db 2878 CGTCGGCAGCTCTCACTTCAGGACAACCTACAGCACATGCTCTCCCCGCCCCAGATCAC 2937
QY 3070 atcgggtccccagagccagctccctcagggccagggagggcagtggtggggcagtggt 3129
Db 2938 ATCGGTCCCCAGAGGCCAGCTCCCTCAGGGCCAGGAGGGGCGAGTGGTGGGGCGAGTGGT 2997
QY 3130 gggggcggtggggccagccacctcccttcagagggggcaaaatctcagcaggttgacagt 3189
 ,
 ,

Db 2998 GGGGGCGGTGGGGGCCAGCCACCTCCCTTTCAGAGGGGGCAAAATCTCAGCAGTTGACAGTG 3057
QY 3190 agtgcgtcccaagaaaacccccggccgtccagcgggaacctattgcagtcctcccgaaaccaagt 3249
Db 3058 AGTGCTGCCAGAAAACCCCGCGCTCCAGCGGGAACCTATTTCAGTCTCCCGGAACCAAGT 3117
QY 3250 tatggtcctgcccgtccacagcgaacagagcctcagcaaaagggcagcatggygcagc 3309
Db 3118 TATGTTCTTCCCCGTCCACGGCAACAGAGCCTCAGCAAAGAGGGCAGCATTTCC 3177
QY 3310 gggggcagcgggtggcgagggggtggggggtcaagccctccatcaccaagcagcatcc 3369
Db 3178 GGGGCAGCGGTGGCGAGGGGTGGGGGCTCAAGCCCTCCATCACCAAGCAGCATTCC 3237
QY 3370 cagactccatccacgctgaacccacgacgtccggcctcgagcgactgtagcctgggtg 3429
Db 3238 CAGACTCCATCCACGCTGAACCCCAAGATGCCGGCTCGGAGCGGACTGTAGCCTGGGTG 3297
QY 3430 tccaatatgcctcacctgtccgctgacatcgagagtgcacacattgagcgggaagagtac 3489
Db 3298 TCCAAATATGCTTACCTGTCCCGTGCACATCGAGAGTGCACACATTTGAGCGGGAAGATAC 3357
QY 3490 aagctgaaggagtagtactcgaagtccatggacgagagccgactggacaggggtgaaggagtac 3549
Db 3358 AAGCTGAAGGAGTACTCGAAGTCCATGGACGAGAGCCGACTGGACAGGGTGAAGGAGTAC 3417
QY 3550 gaggagggagatccactcactgaaggaaaaggctacacatgttccaaacccggaagctggaagag 3609
Db 3418 GAGAGGAGATCCACTCACTGAAGAAAGGCTACACATGTCCAACCCGGAAGCTGGAAGAG 3477
QY 3610 taagagcggaggtgctgtccaggaagagcagaccagcaagatcctgatgcagtaccaa 3669
Db 3478 TAGAGCGGAGGTGTGTCTCCAGGAAGAGCAGACCAGAGATCCTGTATGCAGTACCAA 3537
QY 3670 gccgcctggagcagagcaggaagcgttgaggcagcagcagcagcagcagcagcagcagcag 3729
Db 3538 GCCGCTTGAGCAGAGCGAGAGCGCTTGAGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3597
QY 3730 atcaagagcatcattggcagcgtgatgctggtggaggagcgtgcgcgggagaccacccc 3789
Db 3598 ATCAAGAGCATCATTTGGCAGGCTGATGTGGTGGAGGAGGAGCTGCGCCGGGACCCACC 3657
QY 3790 gccatggctgagccgctgcctgaaccccaagcgtgtgacgctggccccaccccttggaaacgctgg 3849
Db 3658 GCCATGGCTGAGCCGCTGCTGAACCCCAAGAAAGAGGCTGCTCGACGCTCAGAGAGGCAGC 3717
QY 3850 ttcccccttgggtccaaacccccgcgtgtgacgctggccccaccccttggaaacgctgg 3909
Db 3718 TTCCCCCTTGGGTCCAACAAACCCCGCTGTACGCTGGCCCCACCTTTGGAACGGCCTGG 3777
QY 3910 cccccccagccccaccc 3969
Db 3778 CCCCCCAGCCCCACC 3837
QY 3970 acaccgcagaccactagccacccagcatcacagacctccttccctgtgcacctacc 4029
Db 3838 ACACCGCAGACCCTAGCCACCCAGCATCACAGACCTCTTCCCTGTGCACCCCTACCCC 3897
QY 4030 gggccacccagcgtcacagacctccttcccagtgcaacccgacccctggaaacatcaccaacc 4089
Db 3898 GGCCCAACCCAGCGTCAAGACCTCTTCCCAGTGCACCCGACCCCGACCCCTGGAACATCACCAACC 3957
QY 4090 accaggactggacgtcaccaaggacagcgggtgtctctcccttaacgcctccttggggc 4149
Db 3958 ACCAGGACTGGACGTCAACAGGGACAGCGGGATGTCTCCCTTAACGCCCTCTCTGGGGC 4017
QY 4150 acctatctcaacccccactgtctccattccagaggggagagtgggaccctcagctgccct 4209
Db 4018 ACCCATCTGTCAACCCCCACTGTCTCCATTCCAGAGGGGAGAGTGGGACCCCTCAGTGCCCT 4077
QY 4210 ctacccccaggaaccacctacccccacacagacccttcaactctggtggtgctatccccat 4269
Db 4078 CTCACCCCGAGGACACCACTTACCCCAACACAGACCCCTTCACCTCTGGGTGCTATCCCCCAT 4137

QY 616 gaggtaacaacatcgtctcgtggacaaaaatgttttgcctgtcgtcgtcgtcgcagccgaaggac 675
Db 418 GAGGTAAACAACATCGTCTGGGACAAAAATGTTTTGCCCTGTCTGGTCTGCAGCCGAAAGGGAC 477
QY 676 aaatggattgagaatctacagaggcgtgtgaaacccaacaaggacaaacagccgcgggta 735
Db 478 AAATGGATTGAGAAATCTACAGAGGGCTGTGAAACCCAAACAAGGACAACAGCCGCCGGTA 537
QY 736 gataacgtgctgaaactatggatcatagaagctcgagagctgccccccaagaagcgatat 795
Db 538 GATAACGTGCTGAAACTATGGATCATAGAAGCTCGAGAGCTGCCCCCCRAAGCGATAT 597
QY 796 tactgcgagttatgcctggaacgacatgctctatgcacggaccacttccaagccccgcctca 855
Db 598 TACTGCGAGTTATGCCCTGGACGACATGCTCTATGCACGGACCACTTTCCAAGCCCCCGCTCA 657
QY 856 gcctcaggagacactgtcttttggggcgagcacttcgagttttaacaacctgcctgctgc 915
Db 658 GCCTCAGGAGACACTGTCTTTTGGGGCGAGCACTTCGAGTTTAAACAACGTGCCGTGCTGTC 717
QY 916 cgggcgcgtgcgctgcacatcgtaccgtgactcggacaaaaaagcggaaggaaggaagga 975
Db 718 CGGGCGCTGCGGCTGCAFTCTGTACCGTGACTCGACAAAAAAGCGGAAGAGGACAAGGCA 777
QY 976 ggctacgttggcctgtgactgttccagtggtccacctggctgggcgcacttcacagag 1035
Db 778 GGCTACGTTGGCCTGCTGACTGTTCCAGTGGCACCCCTGGCTGGGGCATTCACAGAG 837
QY 1036 cagtggtacccccgtgacccctgccaacaggaagtggggctctcgtggggtatgggctcggg 1095
Db 838 CAGTGGTACCCCGTGACCCCTGCCAACAGGAAGTGGGGCTCTGGGGGTATGGGCTCGGGG 897
QY 1096 ggaggaggggggtcagggggcgcgtcagggggcaaaagggaaggaggtcctcgtgtg 1155
Db 898 GGAGGAGGGGGTCAAGGGGGCGGCTCAGGGGGCAAAAGGAAAGGAGGCTGTCTGCTGTG 957
QY 1156 cggctgaaggcccggttacagacaaatgagtatcctgcctgccccaggtgacgtatataaggagttt 1215
Db 958 CGGCTGAAGGCCGTTACCAGACAAATGAGTATCCTGCCCATGGAGCTATATAAGGAGTTT 1017
QY 1216 gcagaaatatgtgaccaaacactaccgcacgtcgtgtgcgctgtcgtgagccccccctcaat 1275
Db 1018 GCAGAAATGTGACCAAACCACTACCGCATGCTGTGTGCCGTGCTGGAGCCCCCCTCAAT 1077
QY 1276 gtcaagggcaaggagggtcgctagtgactggttcacatcctgcctgcaaaagcacaggcaag 1335
Db 1078 GTCAAGGGCAAGGAGGAGGTGCTAGTGCACCTGGTTTACATCCTGCAAAAGCACAGGGCAAG 1137
QY 1336 gccaaaggacttcctttcagacatggccatgtcagaggttagacgggttcatggagcgggaa 1395
Db 1138 GCCAAGGACTTCCTTTAGACATGGCCATGTTCAGAGGTAGACCGGTTTATGGAGCGGGAA 1197
QY 1396 cacctcatattccgcgagaacacgcctgcaccaataaaagccatagaagagtatatgagactg 1455
Db 1198 CACCTCATATTCCGCGAGAACACGCTCGCCACTAAAGCCATAGAAGAGTATATGAGACTG 1257
QY 1456 attggccagaaataacctcaaggatgccattggggagttcactcggggtcctgtgtatgaatct 1515
Db 1258 ATTGGCCAGAAATACCTCAAGGATGCCATTGGGGAGTTTCATCCGGGCTCTGTATGAATCT 1317
QY 1516 gaggagaactgtgaagtagacccccatcaagtgcacagcgtccagctcctgagtgagagaccag 1575
Db 1318 GAGGAGAACTGTGAAGTAGACCCCATCAAGTGCACAGCGTCCAGTCTTGGCAGAGCACCAG 1377
QY 1576 gccaacctcgggatgtgctgtgagtggccctgtgcaaggtggtcaactccccattgcgtg 1635
Db 1378 GCCAACCTGCGGATGTGCTGTGAGTTGGCCCTGTGCAAGGTGGTCAACTCCCATTTGCGTG 1437
QY 1636 ttccccaggaggaggtgaaggaggtgttgcatcatggcggtgcgctgtgcagagcggggc 1695
Db 1438 TTCCCGAGGGAGCTGAAGGAGGTGTTTGCATCATTGGCGGCTGCGCTGTGCAGAGCGGGGC 1497
QY 1696 cgggaggacattgctgacaggctgatcagcgccctgcctcttccctgcgttccctcctgccccg 1755

Db 1498 CGGGAGGACATTTGCTGACAGGCTGATCAGCGCCTCGCTCTCCTGCGTTTCCCTCTGCCCCG 1557
QY 1756 gccatcatgtcggccccagtctgtttggactgatgcaggagtaccacagatgagcagacctca 1815
Db 1558 GCCATCATGTGCCCCAGTCTGTTGGACTGATGCAGGAGTACCCAGATGAGCAGACCTCA 1617
QY 1816 cgaacctcacctcatcgcgccaaggttatccagaacctcggcacaacttttccaagtttacc 1875
Db 1618 CGAACCTTCACCTTCATCGCCRAAGGTTATCCAGAACCTTGCCAACTTTTCCAAGTTTACC 1677
QY 1876 tcaaggaggagacttcctcctggggttcaatgaacgaggtttctggagctggagtggggttctatg 1935
Db 1678 TCAAGGAGGACTTCTCTGGGCTTCATGAACGAGTTTCTGGAGCTGGAATGGGGTTCTATG 1737
QY 1936 cagcaattcttgatatgagatatccaaacctggacacactgacaaacagcagcagttttgag 1995
Db 1738 CAGCAATTCTTGATGAGATATCCAACCTTGGACACACTGACCAACAGCAGCAGTTTGTAG 1797
QY 1996 ggctacatatagacttgggcgcgagctctccacacttcacacctgcctcgtcgtgggaggtgctg 2055
Db 1798 GGCTACATAGACTTGGGCGCGAGCTCTCCACACTTCACGCCCTGCTCTGGGAGGTGCTG 1857
QY 2056 cccagctcagcaaggaaagccctcctgaagctgggcccgcctgccccgcctcctcctcagcgac 2115
Db 1858 CCCAGCTCAGCAAGGAAAGCCCTCCTGAAGCTGGGCCCGCTGCCCGGCTCTCTCAGCGAC 1917
QY 2116 atcagcacagccctgaggaaccctaaacatccaaaggcagccgagccgcagagcagagcgc 2175
Db 1918 ATCAGCACAGCCCTGAGGAACCCCTAACATCCAAAGGCAGCCGAGCCGCCAGAGCGGCGC 1977
QY 2176 gctcgggtcctcagccccatggtgctgcgcgggcgtcagccgagatgcagggtacacatgatg 2235
Db 1978 GCTCGGTCTCAGCCCCATGGTGTCTGCGGGCCGTACAGCCGTAAGCCGAAATGCAGGGCTACATGATG 2037
QY 2236 cgggacctcaacagctccatcgaccttcagcttccttcagtgcgcagggtcaggtcacaacgtct 2295
Db 2038 CGGGACCTCAACAGCTCCATCGACCTTCAGTCTTCATGGCTCGAGGCCCTCAACAGCTCT 2097
QY 2296 atggacatggctcgcctccctcccccaacaaaggagaaaaacccccgcgcgcctccccggt 2355
Db 2098 ATGGACATGGTCTGCTCCCTCCCTCCCAACCAAGGAGAAACCCCGCGCCCTCCCGGT 2157
QY 2356 gggggtaaagacctgttctatgtgagccggccaccacctggcccggctcctccccagcatac 2415
Db 2158 GGGGGTAAAGACCTGTTCTATGTGAGCGCGGCACCACTGGCCCGGTCTCTCCCCAGCATAC 2217
QY 2416 tgacagagcagctcggacatcacagagccggagcagaagatgctgagtgtcaacaagagt 2475
Db 2218 TGCACGAGCAGCTCGGACATCACAGAGCCCGGAGCAGAAAGATGCTGAGTGTCAACAAGAGT 2277
QY 2476 gtgtccatgctggacctgcagggcgacgggcctggggccttaacagcagtagtgtt 2535
Db 2278 GTGTCCATGCTGGACCTGCAGGGGCGACGGGCTGGGGCGCCCTTAAACGACAGTAGTGT 2337
QY 2536 tccaaacctggcagctgttgggacctgttgcaactcaagccagggttcaactgacagcagcc 2595
Db 2338 TCCAACCTGGCAGCTGTTGGGACCTGTTCACACTCAAGCCAGGCTTCACTGACAGCAGCC 2397
QY 2596 ttgggggttgcgcctgcacctgcgggcctctcccaaggagtggtccttccatcac 2655
Db 2398 TTGGGGTTGGCGCTGCACCTGCGGGCGCCTCTCCCAAGGGAGTGGCTCTTCCATCACA 2457
QY 2656 gcagccggcatgcgcctcagccagatgggtgtcactacggtggtgtccccgccccagcaa 2715
Db 2458 GCAGCCGGCATGCGCCTCAGCCAGATGGGTGTCACTACGGATGCTGTCCCCGCCAGCAA 2517
QY 2716 ctgcgcatccctcttccctccagaacctcttccatataatggtgcggtggaccagg 2775
Db 2518 CTGCGCATCCCTCTTCTCTCCAGAACCCCTCTCTTCCATATGCTGCCGATGGACCAGGG 2577
QY 2776 cccccagcaggccatggaggagcagtggccatggtccaccttccctcccatcaccaccac 2835

Db 2578 CCCCAGCAGGCCATGGAGGGAGCAGTGGCCATGGTCCACCTTCTCTCCCATCACCAACCAC 2637

QY 2836 caccaccatcaccatcacagagggggagaaacccccagggacaccttttgccccgttccat 2895

Db 2638 CACCACCATCACCATACCGAGGGGGAGAACCCCCAGGGGACACTTTTGCCCCGTTCCAT 2697

QY 2896 ggctatagcaagagcgaggacctctctacaggggtccctaagccccctgcgccctccatc 2955

Db 2698 GGCTATAGCAAGAGCGAGGACCTCTCTACAGGGGTCCCTTAAGCCCCCTCGGCCCTCCATC 2757

QY 2956 ctteacagccacagctacagtgatgagtttggacctctgttactgattttaccctcgg 3015

Db 2758 CTTACAGCCACAGCTACAGTGATGAGTTTGGACCCCTCTGGTACTGATTTTACCCGTCGG 2817

QY 3016 cagctctcacttcaggacaacctacagcacatgctctccccggccccagatcaccatcggt 3075

Db 2818 CAGCTCTACTTCAGGACAACCTACAGCACATGCTCTCCCCGCCCCAGATCACCATCGGT 2877

QY 3076 cccagagggccagctccctcagggccagggaggggcagtggtggggcagtggtgggggc 3135

Db 2878 CCCCAGAGCCAGCTCCCTCAGGGCCAGGAGGGGGCAGTGGTGGGGCAGTGGTGGGGC 2937

QY 3136 ggtggggccagccacctcccttgccagagggggcaaaatctcagcagttgacagtgagtct 3195

Db 2938 GGTGGGGCCAGCCACCTCCCTTGCAGAGGGGCCAAATCTCAGCAGTTGCACAGTGAGTCT 2997

QY 3196 gccagaaaaccccgccgtccagcgggaacctattgcagtcctcccggaaccaagtattggt 3255

Db 2998 GCCCAGAAACCCCGCCGCTCCAGCGGGAACCTATTGCAGTCCCCGGAAACCAAGTTATGGT 3057

QY 3256 cctgccccgtccacgggcaacagagcctcagcaagagggcgagcattgggggcagcgggggc 3315

Db 3058 CCTGCCGCTCACGGCAACAGAGCCTCAGCAAGAGGGCAGCATTTGGGGCAGCGGGGC 3117

QY 3316 agcgggtggcggaggggtggggggtcgaagcctccatcccaagcagcattccccagact 3375

Db 3118 AGCGGTGGCGAGGGGTGGGGGCTCAAGCCCTCCATCACCAAGCAGCATTTCCAGACT 3177

QY 3376 ccatccacgctgaacccccacgatgccggcctcggagcggactgtagcctgggtgtccaat 3435

Db 3178 CCATCCACGCTGAACCCACGATGCCGSCCTCGGAGCGGACTGTAGCCTGGGTGTCCAAT 3237

QY 3436 atgcctcacctgtccgtgacatcgagagtgacacattgagcgggaagagtacaaagctg 3495

Db 3238 ATGCTCACCTGTCCGCTGACATCGAGAGTGCACACATTGAGCGGGAAGAGTACAAGCTG 3297

QY 3496 aaggagtaactcgaagtccatggacgagagccgactggacaggggtgaaggagtacgaggag 3555

Db 3298 AAGGAGTACTCGAAGTCCATGGACGAGAGCCGACTGGACAGGGTGAAGGAGTACGAGGAG 3357

QY 3556 gagatccactcactgaaggaaaggctacacatgtccaaacgggaagctggaagagtacagag 3615

Db 3358 GAGATCCACTCAGTGAAGGAAAGGCTACACATGTCCAACCCGGAAGCTGGAAGAGTACGAG 3417

QY 3616 cggaggctgctgtcccaggaagagcagaccagcaagatcctgatgcagtaccacgcccgc 3675

Db 3418 CGGAGGTGCTGTCCCAGGAAGAGCAGACCAAGAGATCCTGATGCAGTACCAAGCCCGC 3477

QY 3676 ctggagcagagcgaagcgcttgaggcagcagcaggtggagaaggactcccagatcaag 3735

Db 3478 CTGGAGCAGAGCGAGAGCGCTTGAGGCAGCAGCAGGTGGAGAAGGACTCCCAGATCAAG 3537

QY 3736 agcatcattggcaggctgatgctggtggaggaggagctgcgcgggaccaccccgccatg 3795

Db 3538 AGCATCATTTGGCAGGCTGATGCTGGTGGAGGAGAGCTGCGCCGGGACCACCCGCCATG 3597

QY 3796 gctgagccgctgcctgaacccaagaagagaggtgctgcgacgtca ---gagaggcagctt 3851

Db 3598 GCTGAGCCGCTGCCTGAACCCAAAGAAGAGAGGCTGCTCGACGCTCAGGTGGAGAGGCAGCTT 3657

QY 3852 ccccccttggtccaaacaaaccccgctgtgacgtgccccaccttggaaacggcctggcc 3911

Db 3658 CCCCCCTTGGGTCCAAACAAACCCCGCGTGTGACGCTTGGCCTTGGACCTTGGAAACGGCCTGGCC 3717

QY 3912 cccccagccccaccccccccccccggtgcagatcacagagaacggcgagttccggaac 3971

Db 3718 CCCCAGCCCCACCCCCCCCCACCCCGGTGCAGATCACAGAGAAGGGGAGTTCCGGAC 3777

QY 3972 accgcagaccactagccccaccacccagatcacagacacctcttccctgtgcaccctaccgccg 4031

Db 3778 ACCGCAGACCACTAGCCCCACCCAGCATCACAGACCTCTTCCCTGTGCACCTACCCCG 3837

QY 4032 cccaccagcgtcacagacctcttcccagtcaccccgaccccgaccctggaaacatcaccaaac 4091

Db 3838 CCCACCCAGGTCACAGACCTCTTCCCAGTGCACCCGACCCCTGGAACATCACCAACCAC 3897

QY 4092 caggactggcgtcaccaagggacagcgggattgtctcccttaacgcctccttggggcac 4151

Db 3898 CAGGACTGGACGTCACCAAGGGACAGCGGATTGTCTCCCTTAACGCCTCCTTGGGGCAC 3957

QY 4152 ccatctgtcaacccccactgctccattccagggaggaggtgggacctcagctgccctct 4211

Db 3958 CCATCTGTCAACCCCACTGCTCCATTCCAGGAGGGAGAGTGGGACCCCTCAGCTGCCCTCT 4017

QY 4212 caccocaggacacacactacccccacacagaccccccttcactctgggtgctatccccatcc 4271

Db 4018 CACCCAGGACACCACTACCCACACACAGACCCCTTCACTCTGGSGTGCTATCCCATCC 4077

QY 4272 t 4272

Db 4078 T 4078

RESULT 5

AB016962 AB016962 4801 bp mRNA ROD 01-SEP-2000
LOCUS Rattus norvegicus mRNA for synGAP-d, complete cds.

DEFINITION AB016962
ACCESSION AB016962
VERSION AB016962.1 GI:4417206
KEYWORDS synGAP-bl.
SOURCE Rattus norvegicus (sub_species:Sprague Dawley) cDNA to mRNA.

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 4801)
AUTHORS Suzuki,T.
TITLE SynGAP-d
JOURNAL Published Only in DataBase (1999) In press
REFERENCE 2 (bases 1 to 4801)
AUTHORS Suzuki,T.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-1998) to the DDBJ/EMBL/GenBank databases. Tatsuo Suzuki, Shinshu University School of Medicine, Department of Neuroplasticity; 3-1-1 Asahi, Matsumoto, Nagano 390-8621, Japan (E-mail:suzukit@sch.md.shinshu-u.ac.jp, Tel:+81-263-37-2683, Fax:+81-263-37-2725)

COMMENT On Mar 16, 1999 this sequence version replaced gi:4239945.
Sequence updated (09-Feb-1999).

FEATURES Location/Qualifiers
1. .4801

source
/organism="Rattus norvegicus"
/sub_species="Sprague Dawley"
/db_xref="taxon:10116"
1124. .4624
/gene="synGAP-d"
/note="synaptic ras GTPase activating protein"
/codon_start=1
/product="synGAP-d"
/protein_id="BAA74972.1"
/db_xref="GI:4417207"

gene

CDS

/translation="MEYFPAAPFRPSQFLSRRLKSSIKRTKSQPKIDRTSSFRJILP
RFRSADHDRARLMQSFKESHESLLSPSSAAEALELNLEDSILKPVHSSILAQEFC
FEVTTSSGTTKCFACRSAAERDKWIENLQRAVKPNKDNRRVDNVLKIWIENLQPK

QY 2107 ctcagcgacatcagcacagccctgaggaacccctaacaatccaaaaggcagccgagcgccag 2166
|||||
Db 2924 CTCAGCGACATCAGCACAGCCCTGAGGAACCCCTAACATCCAAAGGCAGCGAGCGCCAG 2983

QY 2167 agcgagcgcgctcggtctcagcccatggtgctgcgcgggcccgtcagccgagatgcagggc 2226
|||||
Db 2984 AGCGAGCGCGCTCGGTCTCAGCCGATGGTGCTGCGCGGGCGCTCAGCCGAGATGCAGGGC 3043

QY 2227 tacatgatcggggacctcaacagctcccatgcaccttcagtctcttcattcatggctcgagccctc 2286
|||||
Db 3044 TACATGATCGGGACCTCAACAGCTCCATCGACCTTCAGTCTTCATGGCTCGAGGCTC 3103

QY 2287 aacagcttatggacatggctcgctccctccctccccaaccaaggagaaacccccgcgcgc 2346
|||||
Db 3104 AACAGCTTATGGACATGGCTCGCTCCCTCCCTCCCAACCAAGAGAAACCCCCCGCGCC 3163

QY 2347 cctcccggtggggttaaagacctgttctatgtgagccggccaccactggcccggctctcc 2406
|||||
Db 3164 CCTCCCGGTGGGGTTAAAGACCTGTTCTATGTGAGCCGGCCACCACCTGGCCCGGTCCTCC 3223

QY 2407 ccagcatactgcacgagcagctcgacatcacagagccggagcagagaagtctgagtgc 2466
|||||
Db 3224 CCAGCATACTGCACGAGCAGCTCGGACATCACAGAGCCGGAGCCGAGATGCTGAGTGTC 3283

QY 2467 aacaagagtgttccatgctggacctgcagggcgacgggacctgttgcaactcaagccaggtctcactg 2526
|||||
Db 3284 AACAGAGTGTGTCCATGCTGGACCTCGAGGCGACGGGCTTCCCAAGGGAGTGGCTCT 3343

QY 2527 agtagtgtttccaaacctggcagctgttggggacctgttgcaactcaagccaggtctcactg 2586
|||||
Db 3344 AGTAGTGTTCACACCTGGCAGCTGTTGGGACCTGTTGCACCTCAAGCCAGGCTTCACGTG 3403

QY 2587 acagcagccttggggttgcggcctgcaacctgcgggcgccctctcccaaggagtggtctct 2646
|||||
Db 3404 ACAGCAGCCTTGGGGTTGGGGCTGCACCTGCCGGGGCGCCTCTCCCAAGGGAGTGGCTCT 3463

QY 2647 tccatcacagcagccggcatgcgcctccttctccagaacccctctcttcocatatggctgccc 2706
|||||
Db 3464 TCCATCACAGCAGCGGGCATGCGCCTCAGCCAGATGGTGTCACCTACGGATGGTGTCGCC 3523

QY 2707 gccagcaactgcgcacctctcttctccagaacccctctcttcocatatggctgccc 2766
|||||
Db 3524 GCCAGCAACTGCGCATCCCTCTTTCCCTCCAGAACCCTCTCTTCATATGGCTGCCGAT 3583

QY 2767 ggaccaggggccccagcaggccatggaggagcagtgggccatggtccaccttccctcccat 2826
|||||
Db 3584 GGACCAGGGCCCCAGCAGGCCATGGAGGGAGCAGTGGCCATGGTCCACCTTCTTCCCAT 3643

QY 2827 caccaccaccaaccatcaccatcacccgaggggggagaaacccccaggggacacttttgcc 2886
|||||
Db 3644 CACCACCACCACCACCATCACCATCACCGAGGGGGAGAACCCCCAGGGGACACTTTTGCC 3703

QY 2887 ccgttccatggctatgaaagagcagcaggaacctctctacaggggtccctaagccccctgcg 2946
|||||
Db 3704 CCGTTCCATGGCTATAGCAAGAGCGAGGACCTCTCTACAGGGGTCCCTAAGCCCCCTGCG 3763

QY 2947 gcctccatccttcacagccacagctacagtgatgagtttggacctctggtactgatttt 3006
|||||
Db 3764 GCCTCCATCCTTCACAGCCACAGCTACAGTGATGAGTTTGGACCCCTCTGGTACTGATTTT 3823

QY 3007 acccgctcgagctctcacttcaggacaacctacagcacatgctctcccccccgccagatc 3066
|||||
Db 3824 ACCCGTCGGCAGCTCTCACTTCAGGACAACCTACAGCACATGCTCTCCCCCGCCCGAGATC 3883

QY 3067 accatcggtccccagaggccagctccctcagggccagggaggggcagtggtgggggcagt 3126
|||||
Db 3884 ACCATCGGTCCCCAGAGGCCAGCTCCCTCAGGGCCAGGGGGGCAGTGGTGGGGGCAGT 3943

QY 3127 ggtggggggcggtggggccagccacctccctcttcagagggggcaaatctcagcagttgaca 3186
|||||
Db 3944 GGTGGGGCGGTGGGGCGCAGCCACCTCCCTTTCAGAGGGGGCAAAATCTCAGCAGTTGACA 4003

QY 3187 gtgagtgctgccccagaaaaccccgccgtccagcggggaacctattgagtcctcccggaacca 3246

Db 4004 GTGAGTGTGTCAGAAACCCCGCGCTCCAGCGGGAACCTATTGCAGTCCCCGGAACCA 4063
|||||
QY 3247 agttatggtcctgcccgtccacggcaacagagcctcagcaaaagagggcagcattgggggc 3306
|||||
Db 4064 AGTTATGGTCTGCCCCGTCCACGGCAACAGAGCCTCAGCAAAAGAGGCGCAGCATTTGGGGC 4123

QY 3307 agcgggggcagcggtggcgagggggtgggggggctcaagccctccatcaccaagcagcat 3366
|||||
Db 4124 AGCGGGGGCAGCGGTGGCGAGGGGTGGGGGGCTCAAGCCCTCCATCACCAGCAGCAT 4183

QY 3367 tcccagactccatccacgctgaacccccacgctgcggcctcgagcgagcactgtagccctgg 3426
|||||
Db 4184 TCCCAGACTCCATCCACGCTGAACCCCCACGATGCCGGCTCGGAGCGGACTGTAGCCTGG 4243

QY 3427 gtgtccaatatgcctcacctgtccgctgacatcgagagtgcacacattgagcggggaagag 3486
|||||
Db 4244 GTGTCCAATATGCCTCACCTGTCCGCTGACATCGAGAGTGCACACATTGAGCGGGAAGAG 4303

QY 3487 tacaagctgaaggagtactcgaagtccatggacgagagccgactggacagggtgaaggag 3546
|||||
Db 4304 TACAAGCTGAAGGAGTACTCGAAGTCCATGGACGAGAGCCGACTGGACAG-----GGAG 4357

QY 3547 tacgaggaggagatccactcactgaaggaaaggctacacatgtccaaacccggaagctggaa 3606
|||||
Db 4358 TACGAGGAGGAGATCCACTCACTGAAGGAAAGGCTACACATGTCCAACCGGAAGCTGGAA 4417

QY 3607 gagtacgagcggagctgctgtctcccaggaaagagcagaccagcaagatcctgatgcagtac 3666
|||||
Db 4418 GAGTACGAGCGGAGGCTGCTGTCCCAGGAAGAGCAGACCAGCATCTTGATGCAGTAC 4477

QY 3667 caagccccctggagcagagcgaagcgttgaggcagcagcagcagcagcagcagcagcagc 3726
|||||
Db 4478 CAAGCCCGCTTGGAGCAGAGCGAGAAGCGCTTGAGGCAGCAGCAGGTGGAGAGGACTCC 4537

QY 3727 cagatcaagagcatcattgg-----caggctgatgctggtggaggaggagct 3773
|||||
Db 4538 CAGATCAAGAGCATCATTTGGCAGCCCGTCCCTTTCAGGCTGATGCTGGTGGAGAGGAGCT 4597

QY 3774 gcgcggggaccacccccgcctggtgagcgcgtgcctgaacccaagaagagcgtgctcga 3833
|||||
Db 4598 GCGCGGGACCAACCCCGCATGGCTGAGCCGCTGCTGAACCCCAAGAGAGGCTGCTCGA 4657

QY 3834 cgctca-gagaggcagcttcccccttgggttccaaacaccccgctgtgacgctggcccc 3892
|||||
Db 4658 CGCTCAGGAGAGGCAGCTTCCCCCTTGGGTCCCAACAAACCCCGCTGTGACGCTGGCCCC 4717

QY 3893 accttgaacggcctggccccccagccccaccccccccccccccccccccccccccccccccc 3952
|||||
Db 4718 ACCTTGGAAACGGCCTGGCCCCCCCCCAGCCCCACCCCCCCCCCGGCTGCAGATCACAGA 4777

QY 3953 gaacggcgagttcccggaacacgc 3976
Db 4778 GAACGGCGAGTTCCCGGAACACCGC 4801

RESULT 6
HSDJ570F3/c

LOCUS

DEFINITION
HSDJ570F3 94770 bp DNA
Human DNA sequence from clone RP4-570F3 on chromosome 6 Contains a
gene similar to Rattus norvegicus synaptic ras GTPase-activating
protein p135, the C1CK0721Q.5 (polypeptide from patented cDNA
Em:R06811) gene, the PHF1 (PHD finger protein 1) gene, the KNSL2
(kinesin-like 2) gene, the gene for acyl-protein thioesterase,
ESTs, STSs, GSSs and a CpG Island, complete sequence.

ACCESSION
AL050332

VERSION
GI:6010176

KEYWORDS
HTG; c1CK0721Q.5; CpG Island; finger protein; GTPase-activating
protein; kinesin-like; KNSL2; p135; PHD; PHF1; ribosomal protein
L12; RPL12; SynGAP.

SOURCE
human.

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 94770)
Mashreghi-Mohammadi,M.
Direct Submission
Submitted (17-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
On Oct 4, 1999 this sequence version replaced gi:5870478.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

COMMENT

This sequence was generated from part of bacterial clone contigs of
human Chromosome 6, constructed by the Sanger Centre Chromosome 6
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep RP4-570F3 is
from the library RPCI-4 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RP4-570F3 It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.
The true left end of clone RP4-570F3 is at 1 in this sequence. The
true left end of clone ICRF6c-CB2046 is at 94671 in this sequence.
The true left end of clone ICRF6c-CK0721Q is at 20826 in this
sequence. The true right end of clone ICRF6c-CK0721Q is at 61562 in
this sequence.

FEATURES
source

Location/Qualifiers
1. .94770
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone_lib="RPCI-4"
/clone="RP4-570F3"
complement(join(1012..1185,1669..2740,3135..3176,
3870..4048,4255..4456,4664..4900,6710..6854,7066..7210,
7338..7961,10016..10114,10324..10477,12824..12945,
13378..13469,19727..19832,22032..22153,25299..25366))
/gene="dJ570F3.1"
/note="match: cDNAs: Em:AF048976 Em:AF058789 Em:AF058790
Em:AB016962 Em:AF050183 Em:AF047711 Em:AF053938"
/evidence=not_experimental
/product="dJ570F3.1 (homolog of the rat synaptic ras
GTPase-activating protein p135 SynGAP)"
complement(1012..25366)
/gene="dJ570F3.1"
complement(join(<1012..1185,1669..2740,3135..3176,
3870..4048,4255..4456,4664..4900,6710..6854,7066..7210,
7338..7961,10016..10114,10324..10477,12824..12945,
13378..13469,19727..19832,22032..22153,25299..25320))
/gene="dJ570F3.1"
/note="match: proteins: Tr:O95174 Tr:O44242 Wp:CE23473
Wp:CE23471 Wp:CE23475 Wp:CE23472 Wp:CE23470 Wp:CE23474
Wp:CE23476"
/codon_start=1
/evidence=not_experimental
/product="dJ570F3.1 (homolog of the rat synaptic ras
GTPase-activating protein p135 SynGAP)"

mRNA

gene

CDS

/protein_id="CAB63778.1"
/db_xref="GI:6624587"
/translation="MSYAPFDVVRGPGSMHRTQYVHSPYDRPWNPRFCIISGNQLML
DEDEIHPLLIRDRRESSRNKLLRRTVSPVVEGRPHGHEYHLGRSRRKSVPGGKQYS
MEGAPAAPFRPSQGFSLRRLKSSIKRTKSQPKLDRTSFRQILPRFRSADHDRALMQ
SFKESHSHESLLSPSSAAEALNLDSDSIKPVHSSILGQEFCFEVTTSSGTCKFAC
RSAERDKWIENLQRAVKNKNSRRVDNLKLIIEARELPPKKRYICELCDMDMLY
ARTTSKPRSASGDTVFWGEHFEFNNLPVAVRALRLHLVRSDDKKRKDKAGYVGLTVTP
VATLAGRHFEQWYPVTLPTSGSGSGMGSGGGGGGSGGKGGCPAVRLKARYQ
TMSILPMELYKEFAEYVNHVYRMLCAVLEPALNVKKEEVASALVHILOSTGKAKDFL
SDMAMSEVDFRMEREHLIPRENTLATKAIEEYMRLLGQYKLDKDAIGEFIRALYESEN
CEVDPIKCTASSLAHQANLRMCCELALCKVNVNSHCVFRELKEVFASRWLRCAERGR
EDIADRLISASFLFLCPAIMSPSLFGLMQEYPPDEQTSRTLTLIAKVIQNLANFSKF
TSKEDFLGFMNEFLELWGSQQFLYEISNLDTLTNSSSFEgyIDLGRELSTLHALLW
EVLQLSKEALLKGLPLRLINDISTALRNPNIQRPSPRQSERPRQPVVLRGPSAEM
QGYMMRDNLSSIDLQSFMARGLNSMMDMARLPSPTEKPPPPPGGGKDLFYVSRPPL
ARSSPAYCTSSDITEPEQKMLSVNKSVMLDLOGDGGGRLNSSSVSNLAAVGDLLH
SSQASLTAALGLRPAPAGRLSQSGSSITAAGMRLSQMGTVDGVPAQQLRIPLSFQN
PLFHMAADGPGPPGGGGGGGGHPPSSHHHHHHHHHRRGEGPPGDTFAPFHGYSKSED
LSSGVPKPPAAASILHSHYSDEFGPSGDTFRRQLSLQDNLQHMLSPPQITIGQORPA
PSGPGSGSGSGGGGGQPPPLQRGKSQQLTVSAAQKPRPSSGNLLQSPSPSYGPAPR
PRQQSLKEGSIIGSGGGGGGGGLKPSITKQHSQTPSTLNPTMPASERTVAWVSNM
PHLSADIESAHIEREEYKLKEYSKSMDESRLDR"
2178..2210
/note="11 copies 3 mer ggt 90 conserved"
2945..3122
/note="89 copies 2 mer gg 56 conserved"
complement(4211..4599)
/gene="dJ570F3.1"
/note="match: GSS: Em:B68206"
4484..4539
/note="28 copies 2 mer ca 94 conserved"
4605..4649
/note="3 copies 15 mer 91 conserved"
4903..4970
/note="34 copies 2 mer ga 70 conserved"
4996..5190
/note="13 copies 15 mer 59 conserved"
complement(6393..6846)
/gene="dJ570F3.1"
/note="match: GSS: Em:AQ475907"
8077..8178
/note="MLTIC repeat: matches 6..109 of consensus"
8476..8841
/note="MLTIC repeat: matches 109..461 of consensus"
10989..11119
/note="FLAM_C repeat: matches 3..131 of consensus"
13130..13161
/note="16 copies 2 mer ca 93 conserved"
13256..13303
/note="24 copies 2 mer ct 75 conserved"
17111..17815
/note="CpG island"
/evidence=not_experimental
17369..17458
/note="6 copies 15 mer 70 conserved"
17370..17513
/note="72 copies 2 mer cc 63 conserved"
17515..17658
/note="72 copies 2 mer gg 59 conserved"
17631..17726
/note="32 copies 3 mer gcc 63 conserved"
19005..19078
/note="37 copies 2 mer ga 67 conserved"
19612..19691
/note="40 copies 2 mer cc 63 conserved"
19858..19911
/note="27 copies 2 mer cc 70 conserved"
24653..24808
/note="78 copies 2 mer cc 57 conserved"
24706..24831
/note="14 copies 9 mer ctctccct 61 conserved"
25104..25289

repeat_region /note="93 copies 2 mer cc 59 conserved" 25187. .25294

repeat_region /note="12 copies 9 mer ctcccccc 64 conserved" 25367. .25465

repeat_region /note="11 copies 9 mer gggaggag 69 conserved" 25373. .25522

repeat_region /note="10 copies 15 mer 64 conserved" 25413. .25469

misc_feature /note="19 copies 3 mer agg 86 conserved" 26820. .27191

misc_feature /note="match: GSS: Em:AQ605113" 27145. .27728

misc_feature /note="CpG island" 27145. .27728

misc_feature /evidence=not_experimental complement(27412. .27910)

misc_feature /note="match: GSS: Em:AQ369330" 27910. .28149

mRNA join(27713. .28149,28320. .28379,28488. .28533,28661. .28710,28854. .29088)

gene="CICK0721Q.5"

match: cDNAs: Em:AF106943 Em:E06811

match: ESTs: Em:W75983 Em:AA187766 Em:AI004315 Em:R77008 Em:N94792 Em:AA843316 Em:AA384803 Em:AI245618 Em:AI921804 Em:AI792323 Em:AI802586 Em:AI524460 Em:AA843230 Em:AA887246 Em:AI096824 Em:AI929640 Em:N32014 Em:AA725127 Em:AA357305 Em:AA218432 Em:AI332801 Em:AA090206 Em:AI050991 Em:AF034177 Em:W39055 Em:AA622278 Em:AA069440 Em:AI087362 Em:AI168128 Em:AI340323 Em:AI028045 Em:AA934024 Em:AI073391 Em:AA331045 Em:AA708233 Em:AI342398 Em:AA056366 Em:AI582796 Em:AI762297 Em:AI340198 Em:AI541387 Em:N34313 Em:AI567726 Em:AI364919 Em:AI310946 Em:AA181976 Em:AI042256 Em:R81592 Em:T49588 Em:AI929571 Em:D52353 Em:AA035663 Em:N48534 Em:AI373814 Em:AA971499 Em:H25294 Em:AA703121 Em:AA446226 Em:AA379516 Em:AI313425 Em:W70287 Em:AA937820 Em:H21228 Em:AA471384 Em:N71594 Em:AA095718 Em:AA377858 Em:H25235 Em:H52756 Em:D52710 Em:AA329903"

/evidence=not_experimental

/product="dJ570F3.2 (POLYPEPTIDE FROM PATENTED CDNA EMBL:E06811)"

Query Match 21.6%; Score 920.8; DB 69; Length 94770;

Best Local Similarity 91.0%; Pred. No. 7.8e-168;

Matches 979; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 2288 acagctctatggacatggctgcctccctccctcccaaccaaggagaaaccccccgcccc 2347

Db 2744 ACAGCTCTATGGACATGGCTCGCCTCCCTCCCTCCCAACCAAGGAAAGCCACCCACCCAC 2685

QY 2348 ctcccggtggggtaaagacctgttctatatgtgagcggccaccactgccccgtcctccc 2407

Db 2684 CGCCTGGTGGTGAAGACCTGTTCTATGTAAAGCCGTCCACCCCTGGCCCCGTTCCTCAC 2625

QY 2408 cagcactactgcacgacgactcggacatcacagacgagccggagcagaagatgctgagtgtca 2467

Db 2624 CAGCATACTGCACGACGAGCTCGGACATCACAGAGCCAGAGCAGAGATGCTGAGTGTCA 2565

QY 2468 acaagagtgtgtccatgctggacctgcagggcgacggcgccctggggcgcccttaacagca 2527

Db 2564 ACAAGAGTGTGCCATGCTGGACTTACAGGGTGATGGCCCTGGTGGCCGCTCAACAGCA 2505

QY 2528 gtagtgtttccaaacctggcagctgtttggggacctgttgacctcaagccaggcttcactga 2587

Db 2504 GCAGTGTTTCGAACCTTGGCGCCGTAGGGGACCTGCTGCACTCAAGCCAGGCCCTCGCTGA 2445

QY 2588 cagcagccttgggtgcggcctgcacctgcacccgtgcgggcgcctctcccaaggagtggtcttt 2647

Db 2444 CAGCAGCCTTGGGGCTACGGCCTTGGCCTTGGCCGACGCCTCTCCACGGGGAGTGGCTCAT 2385

QY 2648 ccatcacagcagccggcatgctgcctcagccagatgggtgtcactacggatggtgtccccc 2707

Db 2384 CCATCAGGGCGGTGGCATGCGCCTCAGCCAGATGGGTGTCAACCACAGACGGGTGTCCTGT 2325

QY 2708 ccagcaactgcgcacccctcttcccttccagaaacccctctcttccatatggtgcgatg 2767

Db 2324 CCCAGCAACTGCGAATCCCCCTCTCTTCCAGAACCCCTCTCTTCCACATGGCTGCTGATG 2265

QY 2768 gaccagggccccagcaggccatggaggaggagcagtgcccatggtccaccttcctcccatc 2827

Db 2264 GGCCAGGTCCCCAGGGGCCATGGAGGGGGCGGTGGCCATGGCCCCACCTTCTCTCCCATC 2205

QY 2828 accaccaccaccaccatcaccatcacccaggggggagaaacccccaggggagacacttttgccc 2887

Db 2204 ACCACCACCACCACCATCACCAACCAACCGAGGTGGAGAGCCCCCTGGGGACACCTTTGCCC 2145

QY 2888 cgttccatggctatagcaagagcgaggacacctctctctacaggggtccctaaagccccctgcgg 2947

Db 2144 CATTCCATGGCTATAGCAAGAGTGAGGACCTCTCTTCCGGGGTCCCCAAGCCCCCTGCTG 2085

QY 2948 cctccatccttcacagccacagctacagtcatgagtgttgacccctctggtactgattta 3007

Db 2084 CCTCCATCCTTCATAGCCACAGCTACAGTACAGTATGAGTTTGGACCCCTCTGGCAGCTGACTTCA 2025

QY 3008 cccgtcggcagctctcacttcaggagcaaacctcacagcacatgctctccccgccccagatca 3067

Db 2024 CCGTGGCAGCTTTCACTCCAGGACAACCTGCAGCACATGCTGTCTCCCTCCCCAGATCA 1965

QY 3068 ccatcgttccccagagccagctccctcagggccagggccagggagggcagtggtggggcagtg 3127

Db 1964 CCATTGGTCCCCAGAGGCCAGCCCCCTCAGGGCTTGGAGGTGGAGCGGTGGGGGCGACGG 1905

QY 3128 gtgggggcggtgggggcccagccacctcccttgacagaggggcaaatctcagcagttgacag 3187

Db 1904 GTGGGGGTGGGGGGCCAGCCGCTCCATTGCAGAGGGGCAAGTCTCAGCAGTTGACAG 1845

QY 3188 tgagtgtgtccccagaaacccccggccgtccagcgggaaacctattgcagtcctcccggaaccaa 3247

Db 1844 TCAGCGCAGCCCCAGAAACCCCGCCATCCAGCGGGAATCTATTGCAGTCCCCAGAGCCAA 1785

QY 3248 gttatggtcctgcccgtccacggcacaacagagcctcagcaagaggggcagcattgggggca 3307

Db 1784 GTTATGCCCCCGCTCCACGGCAACAGAGCCTCAGCAAGGAGGGCAGCATTGGGGGCA 1725

QY 3308 gcgggggcagcggtgcggagggggtgggggggtcccaagccctccatcaccaagcag 3363

Db 1724 GCGGGGCGACGGGTGGCGGAGGGGTGGGGGGCTGAAGCCCTCCATCACCACAGCAG 1669

RESULT 7

AL161903/c

LOCUS AL161903 150956 bp DNA HTG 29-SEP-2000

DEFINITION Homo sapiens chromosome 6 clone RP11-175A4, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces.

ACCESSION AL161903

VERSION AL161903.12 GI:10039702

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 150956)

AUTHORS Tracey,A.

TITLE Direct Submission

JOURNAL Submitted (28-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk requests: clonerequest@sanger.ac.uk

COMMENT On Sep 9, 2000 this sequence version replaced gi:9796876.

----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: bA175A4

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 99% of reads
Chemistry: Dye-primer-amersham; 0% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 145680 bases at least Q40
Consensus quality: 147560 bases at least Q30
Consensus quality: 148600 bases at least Q20
Insert size: 150756; sum-of-contigs
Insert size: 148326; 10.8% error; agarose-
Quality coverage: 6.43x in Q20 bases; sum-of-contigs Quality
coverage: 6.54x in Q20 bases; agarose-

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 54611: contig of 54611 bp in length
* 54612 54711: gap of 100 bp
* 54712 70293: contig of 15582 bp in length
* 70294 70393: gap of 100 bp
* 70394 150956: contig of 80563 bp in length.
FEATURES
source
1. .150956

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-175A4"
/clone_lib="RPCI-11.1"
1. .54611
/note="assembly_fragment:01454
clone_end:17
vector_side:left"
54712. .70293
/note="assembly_fragment:02542
fragment_chain:1"
70394. .150956
/note="assembly_fragment:02844
fragment_chain:1
clone_end:SP6
vector_side:right"

BASE COUNT 39575 a 37381 c 36411 g 37389 t 200 others
ORIGIN

Query Match 21.6%; Score 920.8; DB 70; Length 150956;
Best Local Similarity 91.0%; Pred. No. 6.9e-168;
Matches 979; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 2288 acagctctatggacatggctgcctccctcccccaacaaaggagaaacccccgcgcgcc 2347
|||||
Db 99893 ACAGCTCTATGGACATGGCTCGCCTCCCTCCCAACCAAGGAAAGCCACCCCAACAC 99834
QY 2348 ctcccgggtggggtaaagacctgttctatgtgagccggccaccactggcccgggtcctccc 2407
|||||
Db 99833 CGCCTGGTGGTGGTAAAGACCTGTTCTATGTAAGCCGTCCACCCCTGGCCCGTTCCTCAC 99774
QY 2408 cagcatactgcacgagctcggacatcacagagccggagcagaagatgctgagtgtca 2467
|||||
Db 99773 CAGCATACTGCACGAGCAGCTCGGACATCACAGAGCCAGAGCAGAAGATGCTGAGTGTC 99714
QY 2468 acaagagtgttccatgtggacctgcagggcgacgggcctggggccgccttaacagca 2527
|||||
Db 99713 ACAAGAGTGTGTCCATGCTGGACTTACAGGGTGATGGGCCTGGTGGCCCGCTCAACAGCA 99654
QY 2528 gtagtgtttccaacctggcagctgttggggacctgttgacctgaagccaggttcaactga 2587
|||||
Db 99653 GCAGTGTTTCGAACCTGGCGCCGTAGGGACCTGCTGCACCTCAAGCCAGGCCTCGCTGA 99594
QY 2588 cagcagccttggggtgcggcctgcacctgcggcgccctctcccaagggagtggtctt 2647
|||||

Db 99593 CAGCAGCCTTGGGGCTACGGCCTGCGCCTGCCGACGCCTCTCCCAGGGGAGTGGCTCAT 99534
QY 2648 ccatcacagcagccggcatgcccctcagcccatcagccagatgggtgtcactaaggtggtgtccccg 2707
|||||
Db 99533 CCATCACGGGGCTGGCATGGCCTCAGCCAGATGGGTGTCACCACAGACGGTGTCCCTG 99474
QY 2708 ccagcaactgcgcataccctcttcttccctccagaacctctcttccatatggctgcccgatg 2767
|||||
Db 99473 CCCAGCAACTGCGAATCCCCCTCTCTCCAGAACCCCTCTCTTCCACATGGCTGCTGATG 99414
QY 2768 gaccagggccccagcaggcccatggaggagcagtgcccatggctccacctctcccatc 2827
|||||
Db 99413 GGCCAGGTCCCCAGCGGCCCATGGAGGGGGGGTGGCCATGGCCCCACCTTCCTCCCATC 99354
QY 2828 accaccaccaccaccatcacccatcacccaggggggagaaacccccaggggacacttttgccc 2887
|||||
Db 99353 ACCACCACCACCACCATCACCCACCGAGTGGAGAGCCCCCTGGGGACACCTTTGCC 99294
QY 2888 cgttccatggtatagcaagagcgaggacctctctctacaggggtccctaaagccccctgcgg 2947
|||||
Db 99293 CATTCATGGCTATAGCAAGAGTGAAGGACTCTCTTCCGGGTCCCCAAGCCCCCTGCTG 99234
QY 2948 cctccatccttcacagccacagctacagtgatgagtttgacacctctggtactgatttta 3007
|||||
Db 99233 CCTCCATCCTTCATAGCCACAGCTACAGTGATGAGTTGGACCCCTTGGCACTGACTTCA 99174
QY 3008 cccgtcgcgagctctcacttcaggacaacctacagcacatgctctccccgccccagatca 3067
|||||
Db 99173 CCCGTGGCAGCTTTCACTCCAGGACAACCTGCAGCACATGCTGTCCCTTCCCCAGATCA 99114
QY 3068 ccacgttccccagaggccagctccctcagggccaggagggggcagtggtggggcagtg 3127
|||||
Db 99113 CCATTGGTCCCCAGAGGCCAGCCCCCTCAGGGCCTGGAGGTGGGAGCGGTGGGGCAGCG 99054
QY 3128 gtggggcggtgggggccagccacctcccttgccagggggcacaatctcagcagttgacag 3187
|||||
Db 99053 GTGGGGTGGCGGGGCCAGCCGCTCCATTGCAGAGGGGCAAGTCTCAGCAGTTGACAG 98994
QY 3188 tgaagtgtccccagaaaaaccccgccgtccagcgggaacctattgcagtcctcccgaaaccaa 3247
|||||
Db 98993 TCAGCGCAGCCCAAGAAACCCCGGCCATCCAGCGGGAATCTATTGAGTCCCCAGAGCCAA 98934
QY 3248 gttatggtcctgcccgtccacggcaacagagcctcagcaaaagagggcagcattgggggca 3307
|||||
Db 98933 GTTATGGCCCCCGCCGTCCACCGCAACAGAGCCTCAGCAAGGAGGCGAGCATTTGGGGGCA 98874
QY 3308 gcgggggcagcgggtggcgagggggtgggggggtcaagccctccatcaccaagcag 3363
|||||
Db 98873 GCGGGGCAGCGGTGGCGGAGGGGGTGGGGGCTGAAGCCCTTCATCACCAAGCAG 98818

RESULT 8
AL161906
LOCUS
DEFINITION Homo sapiens chromosome 6 clone RP11-567N9, *** SEQUENCING IN
PROGRESS ***, 24 unordered pieces.
ACCESSION AL161906
VERSION AL161906.5 GI:9863619
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 169323)
AUTHORS Sims,S.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT On Aug 21, 2000 this sequence version replaced gi:9212890.
----- Genome Center
Center: Sanger Centre
Center code: SC

BASE COUNT 41674 a 41332 c 41269 g 42673 t 2375 others
ORIGIN /note="assembly_fragment:01270"

Query Match 21.6%; Score 920.8; DB 70; Length 169323;
Best Local Similarity 91.0%; Pred. No. 6.7e-168;
Matches 979; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 2288 acagctctatgacatggtcgcctccctcccaacaaagagagaaaccccgccgccc 2347
|||||
Db 83616 ACAGCTCTATGACATGGCTCGCCTCCCTCCCAACCAAGGAAACCCACCCAC 83675

QY 2348 ctcccggtggggtaaagacctgttctatgtgagccggccacactggccggtcctcc 2407
|||||
Db 83676 CGCCTGGTGGTAAAGACCTGTTCTATGTAAGCCGCTCCACCCCTGGCCGCTC 83735

QY 2408 cagcactactgcacagcagctcgacatcacagccggagcagaagatgctgagtga 2467
|||||
Db 83736 CAGCATACTGCACGAGCAGCTCGGACATCACAGAGCCAGAGCAGAGATGCTGAGTG 83795

QY 2468 acaagagtgtccatctgacacctgcagggcgacggcctggggccgaccttaacagca 2527
|||||
Db 83796 ACAAGAGTGTGCTGACTTACAGGCTGATGGCCTGTGTGGCCGCTCAACAGCA 83855

QY 2528 gtagtcttccaaacctggcagctgttggggacctgttgacactcagccaggttcaciga 2587
|||||
Db 83856 GCAGTGTTCGAACCTGGCGCGCTAGGGGACCTGCTGACTCAAGCCAGGCTCGCTGA 83915

QY 2588 cagcagccttgggttgcgacctgcacctgcggcgccctctccaaagggagtgctctt 2647
|||||
Db 83916 CAGCAGCCTTGGGGCTACGGCCTGCGCCTGCGGACGCTCTCCAGGGGAGTGGCTCAT 83975

QY 2648 ccatacagcagcggcagctgcgacctcagccagatgggtgtcactacgagtggtgtcccg 2707
|||||
Db 83976 CCATCAGCGCGGCTGGCATCGCGCTCAGCCAGATGGGTGTACACAGAGCGGTGCTCCTG 84035

QY 2708 ccagcaactgcgcatccctctcttcccaaacctctcttccatattgctgctgctg 2767
|||||
Db 84036 CCCAGCAACTGCGAATCCGCTCTCTCTTCCAGAACCTCTCTTCCACATGGCTGCTGATG 84095

QY 2768 gaccagggccccacacagcagcagctgagggagcagtgccatggtccaccttctccctc 2827
|||||
Db 84096 GGCCAGGTCCCCAGCGGCCCATGGAGGGGGGTGGCCATGGCCACCTTCTCTCCCATC 84155

QY 2828 accac 2887
|||||
Db 84156 ACCAC 84215

QY 2888 cgttcacatggtatagcaagcagcagcagcagcagcagcagcagcagcagcagcag 2947
|||||
Db 84216 CATTCACATGGCTATAGCAAGAGTGAAGACCTCTCTTCCGGGGTCCCAAGCCCTTCTG 84275

QY 2948 cctccatccttcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3007
|||||
Db 84276 CCTCATCTCTTACATAGCCACAGCTACAGTGATGAGTTTGGACCTCTTGGACCTTCTG 84335

QY 3008 cccgtcggcagctctcacttcaggacacacacacacacacacacacacacacacac 3067
|||||
Db 84336 CCCGTCGGCAGCTTCTACTCCAGGACAACTGACGACATGCTGCTCTCTCTCTCTCTCT 84395

QY 3068 ccacggtccccagagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3127
|||||
Db 84396 CCAATGGTCCCCAGAGCCAGCCCCCTCAGGGCCTGGAGGTGGAGCGGTGGGGGCGAGCG 84455

QY 3128 gtggggggtgtgggggagccacccctcccttgcagaggggcaaatctcagcagttgacag 3187
|||||
Db 84456 GTGGGGGTGGCGGGGCGAGCCGCTCTCAATGACAGAGGGGCAAGTCTCAGCAGTTGACAG 84515

QY 3188 tgaagtgtgcccagaaaccccgccgctccagcggggaacctattgagtcctcccggaacaa 3247
|||||
Db 84516 TCAGCGCAGCCCCAGAAACCCCGGCCATCCAGCGGGGAATCTATTGACAGTCCCCAGAG 84575

QY 3248 gttatggtcctcccgccacgcaacagagccctcagcaagagggcagcattgggggca 3307
|||||
Db 84576 GTTATGGCCCCCGCTCCACGCAACAGAGCCTCAGCAAGGAGGGCAGCATGGGGCA 84635

QY 3308 gcgggggagcagcgggtgagggggtggggggtcagccctccatccacaaagcag 3363
|||||
Db 84636 GCGGGGCGACCGGTGGCGGAGGGGTGGGGGCTGAAGCCCTCCATCACCACAGCAG 84691

RESULT 9
AF047711 4368 bp mRNA PRI 14-JAN-1999
LOCUS Homo sapiens nGAP mRNA, complete cds.
DEFINITION AF047711
ACCESSION AF047711
VERSION AF047711.1 GI:4105588
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4368)
AUTHORS Noto, S., Maeda, T., Hattori, S., Inazawa, J., Imamura, M., Asaka, M. and Hatakeyama, M.
TITLE A novel human RasGAP-like gene that maps within the prostate cancer susceptibility locus at chromosome 1q25
JOURNAL FEBS Lett. 441 (1), 127-131 (1998)
MEDLINE 99093006
REFERENCE 2 (bases 1 to 4368)
AUTHORS Noto, S., Maeda, T., Hattori, S. and Hatakeyama, M.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-1998) Department of Viral Oncology, The Cancer Institute, Japanese Foundation for Cancer Research, 1-37-1 Kami-Ikebukuro, Toshima-ku, Tokyo, Japan

FEATURES
Location/Qualifiers
1..4368
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1q24"
126..3545
/note="ras GTPase activating protein"
/codon_start=1
/product="nGAP"
/protein_id="AAD04814.1"
/db_xref="GI:4105589"
/translation="MQTEVPAERPRRRSISGTSTSEKPNMDTANTSPKVPVGFES
KRLKSGIKRKSQSKLDNRNTSFRPLSLRSTDDRSRGLPKLKSRSRSHESLSLSPGCTVEC
LDLGRGEPVSVKPLHSSILGQDFEVYLSGKCFSCNSASERDKWMENLRVQPN
KDCRRRAENVLRWIIIEAKDLAPKKYFCELCDDTLFARTTSTKADNIFWGHFEF
FSLPPLHSITVHIYKDVKKKKDKNNYVGLNIPASTVGRQFVEKWPVSTPTPNK
GKTGSPSIRKSRFQITILPMEQYKEFAEFTSNYMLCSVLEPVISVRNKEBELACA
LVHILQSTGRADFLTLVMSEVDRGCEHDVLFRENTIATKSIIEYLKLVQQLYLD
ALGEFIKALYESDENECVDPKSSSELIDHOSNLKMCCELAFCIKIINSYCVFPRLEK
EVFASWKQQLNRGKQDISERLISASLFLRFLCPAIPMSLFLNLMQEYPPDRTSRLT
LIAKVIQNLAFNFAKFGNKEEYMAFMNDFLEHGWGMRFLLEISNPTISNTPGFDGY
IDLGRELSVLHSLWEVVSQDKGENSFLOATVAKLGLPLRVLADITKSLTNPPIQQ
QLRRFTEHNSPNVSGSLGLOKIFEDPTDSDHLKLSQSDNTDSYFRGKTLILLVQ
QASSQMTVSEKDERESSLPNGRSVSLMDLQDTHAAQVEHASVMDVPIRLTGSQSLSI
TQVASIKQLRETQSTPQAPQVRPLHPALNPQGLPLSFQNPVYHLNPTIPAMPKA
SIDSSLENLSTASSRSQSNSEDFKLSGPNSSMEDFKRSTQSEDFSRRTHTVPDRHIP
LALPRQNSTQQAIRKVDQGLGARAKAPSLPHSASLRSTGSMVSVSAALVAEPVQN
GSRSRQSSSSRESVPVKVRAIQRTQQVQSPVDSATMSPVRTAAVNLNNGQYED
VEETEQLNDEAKHAKEYEQETIKLERLVSSRRLEVERLLVQEQMQKLLEYKA
RLEDSEERLRQEQEKDSQMSKSIISRLMAVEEELKKDHAEMQAVIDAKQKIIDAQEK
IVSLDSANRLMSALTQVKERYSMQVRNGISPTNPTKLSITENGFEKNSSC"

BASE COUNT 1292 a 1022 c 1012 g 1042 t
ORIGIN

Query Match 14.7%; Score 627.8; DB 10; Length 4368;
Best Local Similarity 62.4%; Pred. No. 4.4e-111;
Matches 1079; Conservative 0; Mismatches 572; Indels 78; Gaps 3;

QY 343 caaggcttcctgagccggaggtctaaagctccatcaaacgtacaaagtcacaacaccccaaa 402
Db 243 CCAGGATTTCTCAGCAAGCGCTGAAAGGCTCCATCAAGAGGACCAAAACCCAGTCAAAG 302
QY 403 cttgaccggaccagcagcttttcgacagatctcctgcctccgctccgaaagtgcgtgaccatgac 462
Db 303 CTTGACAGAAACACGAGCTTTGCG-----GCTTCCATCCCTTCGACGTACAGATGAC 353
QY 463 cgggcccggctgatgcagagagcttcaaggagttctcaactcccatgagtcctcctgctgagtcoc 522
Db 354 AGGTCTCGTGGGCTGCCTAACTAAAGAGTCAACGTTCCCATGAATCCTTGCTGAGCCCA 413
QY 523 agcagtgctgctgagggccctggagctcaaacctggatgaagactccattatcaagccagta 582
Db 414 TGCAGCAGAGTGAATGCTCTGGATCTTGGTAGAGGGGAACCTGTATCAGTGAACCACTT 473
QY 583 cacagctccatcctcggccagagagttctgcttttgaggttaacaacatcgtctgggacaaaa 642
Db 474 CATAGTACCATCTTGGACAAAGACTTCTGCTTTGAGGTTACCTACTTAAAGTGAAGTAAA 533
QY 643 tgttttgctgctgctgctgagccgcaagaggaagacaaatggattgagaaatctacagagggct 702
Db 534 TGCTTACGCTGTAAATCTGCTCTGAGAGAGACAAAGTGGATGGAAACCTTCGCAGGACA 593
QY 703 gtgaaacccaacaaggaacacagcccggtgtagatacgtgctgaaactatggatcata 762
Db 594 GTTCAACCTAATAAGGACAAATTGCAGGCGAGCTGAAATGTTCTTCTGTTATGGATCAT 653
QY 763 gaagctgagagctgccccccaagaagcgcgtatctactgcgagttatgcctggagcacatg 822
Db 654 GAAGCCAAGGACCTTGCCCTTAAAGAAATATTTCTGCGAACTGTGCCTTGATGATACC 713
QY 823 ctctatgacaggaaccttccaaagcccgctcagcctcagcctcagcagacactgtcttttgggc 882
Db 714 CTCTTTGCTGCTACAAACCAGCAAGACCAA-----AGCAGACAATAATTTCTTGGGGC 764
QY 883 gagcactctgaggtttaacaacctgcctgctgctcggcgctgcgctgcgctgcatctgtaccgt 942
Db 765 GAACATTTTGAATTTCTCAGCCTTCCACCTCTTCATAGTATCACAGTTTACATTTACAAG 824
QY 943 gactcggacaaaaagcggaaaggaaggaagcagcagcagcagcagcagcagcagcagcagc 1002
Db 825 GATGTGGAAAAAAGAAAAAAGGACCAAGAAATAATTATGTAGGGCTAGTCAACATCCCC 884
QY 1003 gtggccacctggctggcgccacttccacagagcagtggttacccctgacccctgccaaca 1062
Db 885 ACTGCCAGTGTGACTGGTCGCCAATTTGTAGAAAAGTGGTATCCAGTGAGTACACCTACA 944
QY 1063 ggaagtggggctcctgggggtatgggctcggggggaggaggggggtcagggggcggtca 1122
Db 945 CCCAA----- 949
QY 1123 gggggcaagggaagaggtctgctcctgctgctgctgagcgtgaagcggcttaccagacaatg 1182
Db 950 -----CAAAGGAAAGACAGGAGGACCTTCTATTCCGATTAATCACGTTTCCAAACTATC 1004
QY 1183 agtatcctgcccctgagctatataagaggtttgagaatatgtgacccaactaccgc 1242
Db 1005 ACCATTCTGCTATGGAGCAATACAAAGAATTTGCAGAATTTGTCAACCAGCAACTACACC 1064
QY 1243 atgctgtgctgctgctgagcccgccctcaatgtcaaggcgaagagaggtcgctagt 1302
Db 1065 ATGCTGTGTCTGCTCTGAGCCAGTAATTAGTGTGAGAAATAAAGAGGAGTGGCTTGT 1124
QY 1303 gcactggttcacatcctgcaagcacagcagcagcagcagcagcagcagcagcagcagcagc 1362
Db 1125 GCCTTAGTGCACATTTCTCAAGTACTGGCAGAGCCCAAGGATTTTCTGCTGACTTGGTG 1184
QY 1363 atgtcagaggttagaccggttcattgagcggggaacacctcatattccgcgagaaacacgctc 1422
Db 1185 ATGCTGAGGTGGATCGTTTGGAGAGCATGATGCTTGTATCTTCAGAGAGAACACTATT 1244
QY 1423 gccactaaagccatagaagagtatatgagactgattggccagaaataacctcaagagatgcc 1482

Db 1245 GCCACCAATCCATTTGAGGAATACCTCAAGTTGGTGGGACAAACAGTATCTTCATGAGCA 1304
QY 1483 attggggagttcatccgggctctgtatgaatctgaggagaactgtgaagttagaccccatc 1542
Db 1305 CTGGGGGAGTTTATCAAAGCTTTGTATGAGTCCGATGAGAACCTGTGAAGTGGATCCAGC 1364
QY 1543 aagtgcacagcgtccagtcctggcagagcaccagggccaaacctgpggatgtgctgtgagttg 1602
Db 1365 AAATGTTTCATCTAGTGAACCTGATAGACCATCAGAGCAACCTGAAATGTGCTGTGAGCTG 1424
QY 1603 gccctgtgcaaggtgggtcaactcccatcctgctgcttcccgaggagctgaaggaggttt 1662
Db 1425 GCTTCTGCAAGATCATCAACTCTTACTGCTGTTTCCCTCGTGAGTGAAGAAGTCTTT 1484
QY 1663 gcatcatggcgctgctgctgagagcgggggcggggagacattgctgacagggctgac 1722
Db 1485 GCATCATGGAAGCAGCAGTGCCTGAACCTGGCAAGCAAGACATCAGCGAGAGGCTCATC 1544
QY 1723 agcgcctcgtctcctgctgctcctcctgctcctgcccggccatcatgtgcccagctgttttga 1782
Db 1545 AGTGCTCATTAATTTCTCCGTTTCTGCTGCTCCAGCCATTAATGCTCCAGTCTTTTCAAC 1604
QY 1783 ctgatgcaggagtagccagatgagcagacctcacgaacctcacacctcaccctcatcgccaaagtt 1842
Db 1605 CTATGCAGGAGTATCTGTATGACCCGCACATCTCGACTCTAATCTTATGCGCAAGSTC 1664
QY 1843 atccagaacctggccaaacttttccaaagtttaccctcaaaaggaggagcttccctggcttcag 1902
Db 1665 ATTCAGAACCTGGCCAACTTTGCCAAGTTTGGTAACAAGAGGAATAACATGCGCATTCATG 1724
QY 1903 aacgagtttctggagctggagtggtgggttctatgcagcaattctgtatgatatcccaac 1962
Db 1725 AATGATTTTGTAGAACATGAATGGGTGGAATGAAGCGCTTTCTTTGGAGATCTCTAAT 1784
QY 1963 ctggacacactgaccaaacagcagcagcttttgagggtctacatagacttggccgcgagctc 2022
Db 1785 CCAGACACCATCTCAAAACACCCAGGCTTTGATGTTACATTTGATCTGGGCGGAGAGCTT 1844
QY 2023 tccacacttcacgcccctgctcctgggaggtgctgcccagctcagcaagg 2071
Db 1845 TCAGTTTTCATTCCTTACTGTGGGAAGTGTTCCTCAACTTGATAAGG 1893

RESULT 10

AF053938

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

AF053938

Rattus norvegicus

829 bp mRNA

pl35 SynGAP mRNA, partial cds.

GI:3213253

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 829)

Chen, H.-J. and Kennedy, M.B.

Identification and cloning of a novel 130 kd protein containing a

ras GTPase-activating domain from the rat forebrain postsynaptic

density

Abstr. - Soc. Neurosci. 23, 1466-1466 (1997)

2 (bases 1 to 829)

Chen, H.-J., Rojas-Soto, M., Oguni, A. and Kennedy, M.B.

A synaptic Ras-GTPase activating protein (p135 SynGAP) inhibited by

Cam kinase II

Neuron 20 (5), 895-904 (1998)

98282016

3 (bases 1 to 829)

Chen, H.-J. and Kennedy, M.B.

Direct Submission

Submitted (16-MAR-1998) Biology, California Institute of

Technology, 1200 E. California Blvd., MC 216-76, Pasadena, CA.

AF053938

Rattus norvegicus

829 bp mRNA

pl35 SynGAP mRNA, partial cds.

GI:3213253

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 829)

Chen, H.-J. and Kennedy, M.B.

Identification and cloning of a novel 130 kd protein containing a

ras GTPase-activating domain from the rat forebrain postsynaptic

density

Abstr. - Soc. Neurosci. 23, 1466-1466 (1997)

2 (bases 1 to 829)

Chen, H.-J., Rojas-Soto, M., Oguni, A. and Kennedy, M.B.

A synaptic Ras-GTPase activating protein (p135 SynGAP) inhibited by

Cam kinase II

Neuron 20 (5), 895-904 (1998)

98282016

3 (bases 1 to 829)

Chen, H.-J. and Kennedy, M.B.

Direct Submission

Submitted (16-MAR-1998) Biology, California Institute of

Technology, 1200 E. California Blvd., MC 216-76, Pasadena, CA.

FEATURES 91125, USA Location/Qualifiers
source 1. .829
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
183. .>829
/note="synaptic ras GTPase-activating protein; N-terminal splice variant; synaptic ras-GAP"
/codon_start=1
/product="p135 SynGAP"
/protein_id="AAC23491.1"
/db_xref="GI:3213254"
/translation="MGLRPPPTPTPSGGSGSLPPPSHRQPLRRRCSSCCFPGEVHLG RSRKSVPGKQYSMEAPAPFRPSQGLSRRLKSSIKRTKSQPKLDRTSFRQILP RERSADHRLMQSFKESSHESLSPSAAEALELNDESLIKPVHSSILGQEFCE FEVTTSSGTCFACRSAAERDKWIENLQRAVKPNKDNRRVDNVLKLIIEARELP"
BASE COUNT 182 a 266 c 218 g 163 t
ORIGIN
Query Match 12.4%; Score 530; DB 74; Length 829;
Best Local Similarity 100.0%; Pred. No. 4.6e-92;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 250 gaatacacttggtcgctcagaggaggaagagtgctcccggggggaaacagtacagcatg 309
Db 300 GAATACCACTTGGGTGCTCGAGGAGGAAGAGTGTCCCGGGGGAAACAGTACAGCATG 359
QY 310 gaagccgccccgctgcgcccctccgcccctcgcaaggcttcctgagccggagggctaaaa 369
Db 360 GAAGCCGCCCGCTGCGCCCTCGGCCCTCGCAAGGCTTCTGAGCCGGAGGCTAAAA 419
QY 370 agctccatcaaacgtacaaaagtacaaaccccaaacctgacggaccagcagcttgcagac 429
Db 420 AGCTCCATCAAAACGTACAAAGTACAAACCCCAAACTTGACCGGACCAAGCTTCGACAG 479
QY 430 atcctgctcgtctccgaagtgtgacctgaccatgaccgggcccggctgatgcagagcttcaag 489
Db 480 ATCCTGCTCGCTTCGGAAGTGTGACCATGACCGGGCCCGGCTGATGCAGAGCTTCAAG 539
QY 490 gagtccactcccatgagtcctcgtgagtcctcagcagcagtgctgctgagccctggagctc 549
Db 540 GAGTCTCACTCCCATGAGTCCCTGCTGAGTCCCAAGCAGTGTCTGAGGCCCTGGAGCTC 599
QY 550 aacctgaggaagactccattatcaagccagctacacagctccatccctgggcccagggttc 609
Db 600 AACCTGGATGAAGACTCCATTATCAAGCCAGTACACAGCTCCATCCTTGGGCCAGGAGTTC 659
QY 610 tgccttgaggttaacaacatcgtctgggacaaaatgttttgcctgctcgtgctcagccgaa 669
Db 660 TGCTTTGAGGTAACAACATCGCTGCGGACAAAATGTTTGCCTGTCTGCTGAGCCGAA 719
QY 670 agggacaaatggattgagaatctacagagggtgtaaaaccccaacaaaggacacagccgc 729
Db 720 AGGGACAAATGGATTGAGAACTTACAGAGGGCTGTGAACCCCAACAAAGGACACAGCCGC 779
QY 730 cgggtagataaacgtgctgaaactatgtagatcatagaagctcagagctgcc 779
Db 780 CGGGTAGATAACGTGCTGAAACTATGATCATAGAAAGCTCGAGAGCTGCC 829
RESULT 11
AK024488
LOCUS AK024488 4287 bp mRNA PRI 29-SEP-2000
DEFINITION Homo sapiens mRNA for FLJ00087 protein, partial cds.
ACCESSION AK024488
VERSION AK024488.1 GI:10440489
KEYWORDS fis (full insert sequence).
SOURCE Homo sapiens adult spleen cDNA to mRNA, clone:as00087.
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 4287)
AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
TITLE The nucleotide sequence of a long cDNA clone isolated from human spleen
JOURNAL Published Only in DataBase (2000) In press
REFERENCE 2 (bases 1 to 4287)
AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-2000) to the DDBJ/EMBL/GenBank databases. Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913, Fax:81-438-52-3914)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert construction; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Kazusa DNA Research Institute.
FEATURES Location/Qualifiers
1. 4287
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="as00087"
/dev_stage="adult"
/tissue_type="spleen"
/note="vector:pBluescriptII SK plus"
1339. .3363
/gene="FLJ00087"
1339. .3363
/gene="FLJ00087"
/note="For this clone, GeneMark analysis triggered an alert for spurious CDS split. The result of GeneMark analysis of the nucleotide sequence of this clone is accessible through http://www.kazusa.or.jp/NEDO. Start codon is not identified."
/codon_start=1
/product="FLJ00087 protein"
/protein_id="BAB15778.1"
/db_xref="GI:10440490"
/translation="GSTLHLRCLLWGANKPKRFPSPSHSLQDNVEREETWLSVWVHEA KGLPRAAGAPGVRAELWLDGALLARTAPRAGPQLFWAERFHEALPPARRLSRLR GLPGSAVIGRVALALELDAPRAPAAGLERWFPLLGAPAGAAALRARIRARRLRVLP ERYKELAEFTLHYARLCGALEPAALPAQAEELAAAMRVLRATGRAQALVTDLGTA LARCGGREALFRLTATKAIDYMKLVADYQLOETLGQVVRRLCASTEDCEVDPSK CPASELPEHQARLNSCEEVFETIIHSYDWFPAELGIVFSSWRACKERGSEVLGPR LVCASFLRLCLLPAILAPSLFGLAPDHPAPGAPARTLTIAKVIONLANRAPFGEKEAY GMNSFLHEHGPMQCFLDQVAMVDVDAAPSGYGGSDALQLAVLHAQLCTIFAELD QTRDTLEPLPTILRAIEEGQVPLVSVPMRLPLPAQVHSSLSAGKPGFLAPRDLPK HTPLISKQSLRSVRSESWARPRPDERPLRRPRPVQRTQSVPRRPARRRQSAGPW PRPKGSLSMGPAPRAPRPTWRTDSASLPRKPSVPWQMDQPDNRNALGTHRPVNKLAE LQCEVAALREORVLSRLVESLSTQIRALTEQQELRQLQLDLSRLRAG"
BASE COUNT 785 a 1529 c 1261 g 712 t
ORIGIN
Query Match 7.9%; Score 336.4; DB 34; Length 4287;
Best Local Similarity 59.2%; Pred. No. 5.5e-55;
Matches 574; Conservative 0; Mismatches 396; Indels 0; Gaps 0;
QY 1148 ctgctgtgctgagggccgttaccagacaaatgagtcctgccccatggagctatata 1207
Db 1769 CAGCGCTGCGGGCGGATTCCGGCGCGTCCCTGCGCGTCTGCGCGGCGGTACA 1828
QY 1208 aggagtttgcaagaatatgtgaccaaccactaccgcctgctgctgctgagcccg 1267
Db 1829 AGGAGCTGCGGGAGTTCCTCACCTTCCACTATGCGCGCTCTGCGGGCCCTGGACCCG 1888
QY 1268 ccctcaatgtcaaggcgaaggaggtcgtagtcgactggttcacatcctgcaagca 1327
Db 1889 CGCTGCTGCGCAGGCCAACAGGAGAGCTGCGCGCAGCCATGGTGGCTGCTGCGGCGCA 1948
QY 1328 caggcaaggccaaggacttcttccagacatggccatgtagaggttagaccggttcattg 1387

Db 1949 CCGGCGGGCGCAGGCGTGGTGACTGACCTGGGCACTGCGGAGCTGGCGCGTGTGGAG 2008
QY 1388 agcggaacacacctcatattccgcgagaaacacgctcgccactaaagccatagaagagtata 1447
Db 2009 GCCGTGAGGCGCTGCTGTTCGGGAAACACATTTGGCCACCAAGCATATCATGAGTACA 2068
QY 1448 tgagactgattggccagaaataacctcaaggtatgcccattggggagttcatccgggctctgt 1507
Db 2069 TGAAGCTCGTGGCACACAGGATTACCTCCAGGAGACCCTGGGACAGGTTGTGCGGCGTCTCT 2128
QY 1508 atgaatctgaggagaaactgtgaagttagaccccccatcaagtgcacagcgtccagttctggcag 1567
Db 2129 GTGCTTCTACTGAGGACTGTGAAGTGGACCCCGACCAAAATGTCCAGGCTCGGAGCTGCCAG 2188
QY 1568 agcaacaggccaaacctgcgagatgtgctgtgagttgtggccctgtgaaaggtgtgtaactccc 1627
Db 2189 AGCACCAGGCCAGACTTCGGAACAGCTGCGAGGAGGCTCTTCGAAACCATATCCATTCTCT 2248
QY 1628 attgctgttccccaggaggagctgaaggaggtgtgttgcacatcatggcgctgcgtgtgcag 1687
Db 2249 ACGACTGTGTTCCCTGCGGAGCTGGGCATCGTGTCTCAAGCTGGCGGAGAACATGTAAG 2308
QY 1688 agcgggccgggaggacatgtctgacaggtgacagcgcctcgctcgtcttctcgtctcc 1747
Db 2309 AACGTGGCTCTGAGGTGCTGGGCCCCCGACTGGTGTGGCCTCCCTCTCTCTGCGGCTCC 2368
QY 1748 tctgcccggccatcatgtgccccagctgtgttggactgatgcaggagttaccacagatgagc 1807
Db 2369 TGTGCCCTGCCATCCTGGCACCCAGCCTCTTTGGTTTGGCACCCAGACCATCCAGCACCCG 2428
QY 1808 agacctcaacaaacctcacctcatcgccaaaggttatccagaaacctggccaaacttttcca 1867
Db 2429 GCCCAGCCGCGCACCTTCACACTGATTGCCAAGGTGATCCAGAACCTCGCCAAACCGTGCCC 2488
QY 1868 agtttaacctcaaaaggaggacttcctggttcctgaacagagtttctgagctggagtggg 1927
Db 2489 CGTTCCGTFGAGAAGGAGGCGCTACATGGGCTTCATGAATAGCTTCTCTGGAGGAACATGGAC 2548
QY 1928 gttctatgcagcaattcttctgtatgatataccaaactggacacactgacccaacagcagca 1987
Db 2549 CAGCCATGCAATGCTTCTTGACACAGGATAGCCATGGTGGATGTGGATGTGCTGCCCGCAGTG 2608
QY 1988 gttttgagggctacatagacttgccgagctgtccacacacttcaacacctgctctggg 2047
Db 2609 GTTACCAGGGCAGTGTGTATCTGGCCCTCCAGTGTAGCTGTCTGCTGATGCCAGCTCTGTA 2668
QY 2048 aggtgctgccccagctcagcaaggaagccctcctgaagctgggccccgtgccccggtccc 2107
Db 2669 CAATTTTGTGCTGAGCTTGACCAGACAAACCCGAGACACCCTTGAACCACTGCCACCATCC 2728
QY 2108 tcagcgacat 2117
Db 2729 TGGGAGCCAT 2738

RESULT 12
AC019800/c
LOCUS AC019800 56366 bp DNA HTG 03-JAN-2000
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC019800
VERSION AC019800.1 GI:6665097
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 56366)
AUTHORS Adams,M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA

COMMENT This sequence was identified as CDM:10210991 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
Location/Qualifiers
1. 56366
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT 16234 a 11924 c 11808 g 16400 t
ORIGIN

Query Match 5.7%; Score 241.8; DB 39; Length 56366;
Best Local Similarity 50.6%; Pred. No. 4.5e-37;
Matches 749; Conservative 0; Mismatches 682; Indels 48; Gaps 5;

QY 667 gaaagggacaaatgattgagaatctacagagggctgtgaaaccccaacaaaggacacagc 726
Db 25786 GAGCGGACCTTTGGATCTACTCGCTGGCGCAAGTCGATCGCTCCGAATGCAGACACACG 25727
QY 727 cgccgggtagataacgtgctgaaactatggtatcatagaaagctcgagagctgcccccaag 786
Db 25726 CGTCGACGGACAACTCGCTGAAGATGTGGGTGTACGAGCGCAAAAATCTGCCGCCCAAG 25667
QY 787 aagcgatattactgcgagttatgcctgggacacatgctctatgcacggaccacttccaag 846
Db 25666 AAGCGTTACTTTTGGCAACTGCAATTTGGACAAAGACGCTGTACGGCGCGGACTTCGGTGAAG 25607
QY 847 ccccgctcagcctcaggagacactgtcttttgggcgagacacttcgagtttaacacactg 906
Db 25606 CTGCAGAC-----GGATCTGCTGTGTTTGGGGGAGCACTTCGATTTCCCCGACATA 25556
QY 907 cctgctgtcgggcgctgcggctgcatctgtaccgtgactcggacaaaagcggaagaag 966
Db 25555 CCGGAGATTAATGTGATCACTGTTAAGCTTTTCCGTGAGGTGACAAAGAAAGAACGCGG 25496
QY 967 gacaaggcaggtacgtgtgctggtgactgttccagtgccacccctggctggcgccac 1026
Db 25495 GACAAATACCAATTTGTGGGATCGGTGAAGATACCCGTGCAGGATGTCACTCCAGATTG 25436
QY 1027 ttcacagagcagtggtacccccgtgacctgccaacaggaaagtggtggttatg 1080
Db 25435 CCCTGCGAACAATGGTATCCCAT-----CTGAGCGACAAAGCGCGGC 25394
QY 1087 ggctcgggggaggagggggtcagggggtcagggggtcaggggcaaaagggagggctgt 1146
Db 25393 GACAGTCTGGCGAGGACATCGGGCGGCGCGGCGCAGTGGGTCCAAAGGACAAAGGACAAATTG 25334
QY 1147 cctgctgtcgggtgaagcccgcttaccagacaaatgagtatcctgccccatggagctatat 1206
Db 25333 CCCACGCTGAGGATCAAGTGTCTGTTTCCAGAGACCCGACATCTGCCCATCAATGTGTAC 25274
QY 1207 aaggagtttcagaaatatgtgaccaaccactaccgcatgctgtgtgctgctggagccc 1266
Db 25273 GGCAACTTTTGGCTACCTCAAGGAGAACTATAAGCGCGTGTGCGAGACCCCTGGAGCCG 25214
QY 1267 gccctcaatgtcaaggcgcaaggaggaggtcgctagtgcactggttcacatcctgcgaagc 1326
Db 25213 GTGATCGGAGTCAAGGCCAAGGAGGAGACATGGACAGGCACCTGCTGCTGATGCACGCA 25154
QY 1327 acaggcaaggccaaaggacttccttcagacatggccatgtcagaggttagaccgggttcag 1386
Db 25153 CAGGGATTGGCGGCGCCTTCTCTACCGATGTGGTGGCCCTCGATCTGCTGCGAGTTGGC 25094
QY 1387 gacgggaacacacctcatattccgcgagaaacacgctcgcctactaaagccatagaagagtat 1446
Db 25093 GATCAGAGG---CTTACGTTTCAGGGGCAACTCCTTTGGCCACCAAGAGCATGGAGGCATT 25037
QY 1447 atgagactgattggccagaaataacctcaaggtatgccattggggaggttcatccgggctctg 1506
Db 25036 CTCAAGCTGACGGCGGAACAGTATCTGCGAGGACACACTATCCGACCAATAAACGAGCTA 2497

QY 1267 gccctcaatgtcaaggccaaggagaggtcgttagtgcactggttcacatcctcgcaaacg 1326
Db 113952 GTGATCGGAGTCAAGGCCAAGGAGGACATTGGACAGGCACCTGGTGTCTGATGCACGCA 114011
QY 1327 acaggcaaggccaaggacttcctttcagacatggtccatgtcagaggtagaccggttcctg 1386
Db 114012 CAGGGATTGGCGGGCGCTTCTTCCACCGATGTGGTGGCCCTCGATCTGCTGGAGTTGGC 114071
QY 1387 gaggcggaacacctcatattccgcgagaaacagctcgccactaaagccatagaagatgat 1446
Db 114072 GATCAGAGG--CTTACGTTACAGGGGCAACTCTTGGCCACCAAGAGCATGGAGGCATTC 114128
QY 1447 atgagactgattggccagaaataacctcaagatgccattggggaggttcacccgggtctg 1506
Db 114129 CTCAAGTGTACGGGGCAAGCATATCTGCAGGACACACTATCCGCACCAATAAACGAGCTA 114188
QY 1507 tatgaatctgaggagaactgtgaagttagaccatcaagtgc-----cagcgctcc 1557
Db 114189 ATTACGTCGGAGAGGACTTGGAGGTGGATCCACCAAGACGAGCGGTTCTGTCGGCGGGT 114248
QY 1558 agtctgcagagcaccagccaacctgcggatgtgctgtgagttggccctgtgcaaggtg 1617
Db 114249 TCGTGTGACGACAGCAGCGCGCGCTTTCGTTGGCGCGGTCCGAGGGCGTGGCAGTGCAATC 114308
QY 1618 gtcaactccatcctgctgttcccgagggaggtgaggtgtttgcatcatggtcggtg 1677
Db 114309 TTGGAATCGCACAAAGCAATTTTCCCGCCAGTTACGAAATTTGCTTTGCGACGTTCCGGGAG 114368
QY 1678 cgtgtgcagagcgggcgaggagacattgtgacaggtgtgacgcctcgtcctcttc 1737
Db 114369 CGCTTGACGACAGTGGCGCGCTCAGGATATGGCGGACAACTGTATCTCGCGGAGCAATTTTC 114428
QY 1738 ctgcgttctctctgcggccatcatgtgcgcaggtgtgttgactgatgcagcaggtac 1797
Db 114429 CTGCGTCTCTGTGCGCGGCATCTGTGCGCGCTCTTCTTCAATATCACCAGCGAATG 114488
QY 1798 ccagatgagcagacctaagaaacctcaccctcatcgccaagttatccagaaacctggcc 1857
Db 114489 CCGTCCGACGGGCCACCCGCAATCTCACACTTGTGGGCCAAGACCTGCAAAACATTTGGCC 114548
QY 1858 aacttttccaaagtttacctcaaaaggaggacttctggtggttcataaacgagtttctgag 1917
Db 114549 AATTTACCCCGCTTCCAGGGCAAGAGAACTTTATGGAGTTTCTCAACGATTTCTCTCGAG 114608
QY 1918 ctggagtggttctatgcagcaattcttctgtatgatataccaaacctggacacactgac- 1976
Db 114609 CAGGAGGCGCGCTCGCATGCAACAGTTTCTGGAGATTATATCCACGCGCGCGGAGCACCCA 114668
QY 1977 -----caacagcagcagttttaggggtctacatagacttggcgcgagctctccaca 2028
Db 114669 GCTCCAGACTCGATCTCGATTTGGCGCGGTACATCGACAGGGCAAAACAGTTGTCCATA 114728
QY 2029 cttaagcctctgctctggaggtgctgccccagctcagcaaggagccctcctgaaagctg 2088
Db 114729 CTACACAGTTTGTCTCAGCGAAAGCCTTGCCCAAGCTGCCGAGGCGCAGGCAGCAGCTG 114788
QY 2089 ggcccgctgccccggtctcctcagcgacatcagcacagcc 2127
Db 114789 GATCCGTTGCAGCATATTCTCGATGAAATCAGCCGAGCC 114827

RESULT 14
AC012161/c
LOCUS AC012161 202741 bp DNA HTG 16-NOV-1999
DEFINITION Drosophila melanogaster chromosome X clone BACR01122 (D1114)
RPCI-98 01.I.22 map 16B-16D strain Y; cn bw sp, *** SEQUENCING IN
PROGRESS ***, 142 unordered pieces.
ACCESSION AC012161
VERSION AC012161.7 GI:6435885
KEYWORDS HTG; HTGS_PHASE1.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

REFERENCE
AUTHORS

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 202741)
Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, K.,
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hinkle, A., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K.,
Kearney, L., Lee, B., Lewis, S., Li, P., Ling, H., Moshrefi, A.R.,
Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B.,
Richards, S., Sethi, H., Swirskas, R.R., Wan, K.H., Webster, D.,
Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G.M.
Sequencing of Drosophila melanogaster

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 202741)
Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,
Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P.,
Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,
Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E.,
Swirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and
Rubin, G.M.
Direct Submission

TITLE
JOURNAL
COMMENT

Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Nov 16, 1999 this sequence version replaced gi:6425635.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdgp@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 142 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 509: contig of 509 bp in length
* 510 589: gap of unknown length
* 590 1201: contig of 612 bp in length
* 1202 1281: gap of unknown length
* 1282 2394: contig of 1113 bp in length
* 2395 2474: gap of unknown length
* 2475 3512: contig of 1038 bp in length
* 3513 3592: gap of unknown length
* 3593 4500: contig of 908 bp in length
* 4501 4580: gap of unknown length
* 4581 5465: contig of 885 bp in length
* 5466 5545: gap of unknown length
* 5546 6253: contig of 708 bp in length
* 6254 6333: gap of unknown length
* 6334 7105: contig of 772 bp in length
* 7106 7185: gap of unknown length
* 7186 8077: contig of 892 bp in length
* 8078 8157: gap of unknown length
* 8158 9047: contig of 890 bp in length
* 9048 9127: gap of unknown length
* 9128 9530: contig of 403 bp in length
* 9531 9610: gap of unknown length
* 9611 10173: contig of 563 bp in length
* 10174 10253: gap of unknown length
* 10254 10841: contig of 588 bp in length
* 10842 10921: gap of unknown length
* 10922 11807: contig of 886 bp in length
* 11808 11887: gap of unknown length
* 11888 13165: contig of 1278 bp in length
* 13166 13245: gap of unknown length
* 13246 14526: contig of 1281 bp in length
* 14527 14606: gap of unknown length

Qy 1027 ttacagagcagtggttaccctgtagccctgcccacaggaagtggggctctctggggtatg 1086
 Db 126098 CCCTGCGAACAATGGTATCCATA-----CTGAGCGCAAGGCGGCG 126057

Qy 1087 ggctgggggagggagggggtcaggggggtcaggggggtcaggggggtcaggggggtcgt 1146
 Db 126056 GACAGTCTGGGAGGAGACATCGGGCGGCGGCGGAGTGGGTCCAAGGACAGGAGCAATTG 125997

Qy 1147 cctgtgtgctgagggcgggtaccagacaaatagtagtctcctgcccagcttatat 1206
 Db 125996 CCCACGCTGAGGATCAAGTGTGCTTCCAGAGCACCGACATCTGCCATCAATGTGTAC 125937

Qy 1207 aaggagttgcagaatatgtgaccaaacactaccgcatgctgtgtgcctgtgagcccc 1266
 Db 125936 GGCAACTTTTGACGTACCTCAAGGAGAACTATAAGCGCGTGTGCGAGACCGCTGGAGCGG 125877

Qy 1267 gccctcaatgtcaaggggcaagaggaggtgcgtagtgactggttcacatcctgcgaagc 1326
 Db 125876 GTGATCGGAGTCAAGGCCAAGGAGGACATTTGGACAGGCACTGGTGTCTGTGATGCACGCA 125817

Qy 1327 acaggcaaggccaaggaacttcctttcagacatggccatgtcagaggttagaccggttcacg 1386
 Db 125816 CAGGGATGGCGGGCGCTTCTTCAACCGATGTGGTGGCCCTCGATCTGCTCGAGTTGGC 125757

Qy 1387 gagggggaacacctcatattccgcgagaaacacgcctgcacactaaagccatagaagatat 1446
 Db 125756 GATCAGAGG---CTTACGTTTACGGGGCAACTCTTGGCCACCAAGAGCATGGAGGCATTC 125700

Qy 1447 atgagactgattggccagaaatacctcaaggaatgcatggggagttcctccgggctctg 1506
 Db 125699 CTCAAGCTGACGGGGCAAGCATATCTGCAGGACACACTATCCGCCACCAATAACAGGCTA 125640

Qy 1507 tatgaactgaggagaactgtgaagttagaccccccatcaagtga-----cagcgctcc 1557
 Db 125639 ATTCACTCGGAGAGGACTCGGAGGTGGATCCCAAGAGAGCGGTTCTCGCGCGGT 125580

Qy 1558 agtcggcagagcaccagggccaacactgcggatgtgtgtgagttggccctgtgcaaggtg 1617
 Db 125579 TCGCTGCAGCGACAGCAGGCGCGCTTGGCTGGCGCGGTCCGAGGGCGGTGGCAGTGCAATC 125520

Qy 1618 gtcaactccattgctgttcccgagggagagctgaaggaggtgtttgcatcatgagcgctcttc 1677
 Db 125519 TTCAATCGCACAAAGCATTTCCCGCGCCAGTTACGAAATGCTTTGCGGACGTTCCGGGAG 125460

Qy 1678 cgtgtgcagagcggggcgaggagacattgtcagaggtgctgacagcgctcgtctcttc 1737
 Db 125459 CGCTGCAGCAGCTGGCGCTCAGGATATGGCCGACAACTGATCTCGGCGAGCATTTTC 125400

Qy 1738 ctgcttctcctctgccccgcacatcatgtcgccccagctgtgttgactgagcagagtagac 1797
 Db 125399 CTGCGCTTCTGTGCGCGGCAATCTGCTGCGCGCTCTTCAATATCACAGCGAAGTCTG 125340

Qy 1798 ccagatgagcagacccacgaacccctcacccctcatcgccaaaggttatccagaacctggcc 1857
 Db 125339 CCGTCCGACGCGGCGCACCCGCAATCTCACACTGTGTGGCCAAAGACCCCTGCAAAACATTGGCC 125280

Qy 1858 aactttccaagtgttaccctcaaaaggaggagacttccctgggcttcatgaacgagttctggag 1917
 Db 125279 AATTACCCCGCTTCCAGGCGCAAGAGAACTTTATGAGATTCTCAACGATTCTCTCGAG 125220

Qy 1918 ctgagtgagggttctatgagcaattcttctgtatgatatccaaacctggacacactgac- 1976
 Db 125219 CAGGAGCGCGCTCGCATGCAACAGTTTCTGGAGATTATATCCACGCGCGCGGAGCACCCA 125160

Qy 1977 -----caacagcagcaggttttgagggtacatagacttggggcgagctctccaca 2028
 Db 125159 GCTCAGACTCGATCTCGATTGGCGCGGCTACATCGACCGAGGCAACAGTTGTCCATA 125100

Qy 2029 cttcacgcccctgctctgagggtgctgccccagctcagcaaggaagccctcctgaagctg 2088
 Db 125099 CTACACAGTTTGTCTCAGCGAAAGCCTTGCCCAAGCTGCGCGGAGGCGGAGGAGGAGCTG 125040

Qy 2089 ggcccgctgccccggtcctcagcgacatcagcacagcc 2127
 Db 125039 GATCCGTGCAGCATATCTCGATGAATCAGCCGAGCC 125001

RESULT 15
 AE003506
 LOCUS AE003506 300994 bp DNA INV 04-OCT-2000
 DEFINITION Drosophila melanogaster genomic scaffold 142000013386053 section 23
 of 30, complete sequence.
 ACCESSION AE003506 AE002593
 VERSION AE003506.1 GI:72933355
 KEYWORDS HTG.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 300994)
 REFERENCE
 AUTHORS
 Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
 Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
 George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
 Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
 Brandon,R.C., Rogers,Y.H., Blazek,R.G., Champe,M., Pfeiffer,B.D.,
 Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
 Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J., Basu,A.,
 Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Beeson,K.Y.,
 Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
 Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
 Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
 Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
 Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
 Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
 Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
 Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrera,S.,
 Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S., Ge
 Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
 Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
 Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
 Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
 Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
 Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
 Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
 McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
 Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
 Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
 Nusskern,D.R., Pacieb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
 Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
 Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
 Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
 Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,C., Turner,R.,
 Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
 Weinstein,G.M., Weissbach,J., Williams,S.M., Woodage,T.,
 Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
 Zaveri,J.S., Zhang,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
 Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
 Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
 The genome sequence of Drosophila melanogaster
 Science 287 (5461), 2185-2195 (2000)
 20196006
 2 (bases 1 to 300994)
 Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
 Direct Submission
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 Location/Qualifiers
 1. 300994
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /chromosome="X"
 complement(join(<16198..16317,16386..16574,16635..16845,
 16915..17053,17129..17411,17526..17607,17676..17975,
 18107..18483))

TITLE
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 mRNA

[illegible]

Search completed: January 18, 2001, 06:46:44
Job time: 18715 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: January 18, 2001, 02:10:16 ; Search time 505.11 Seconds
(without alignments)
3177.191 Million cell updates/sec

Title: US-09-294-298-1
Perfect score: 4272
Sequence: 1 atgtcctatgcccccttcag.....tggggtgctatccccatcct 4272

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 16
1: /SIDS6/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.8	2.2	2647	18	T76739
2	63.6	1.5	1337	20	Z17263
3	61	1.4	795	19	V55830
4	61	1.4	799	19	V55831
5	61	1.4	5452	20	X90923
6	61	1.4	8705	20	Z23778
7	61	1.4	9600	19	V21683
8	61	1.4	10380	20	Z22248
9	61	1.4	10596	14	Q51731
10	61	1.4	10596	17	T40348
11	61	1.4	10596	20	X15650
12	59.6	1.4	1925	20	X90924

13	59.4	1.4	1908	8	N71064	Gene encoding Plas
14	59.2	1.4	1000	21	A02484	Human colon cancer
15	57.8	1.4	49999	20	Z23895	Murine LOBO homolo
16	57.6	1.3	2338	19	V23249	Nephila clavipes s
17	57.6	1.3	2338	21	Z38195	N. clavipes spider
18	56	1.3	2000	8	N71065	Gene encoding plas
19	56	1.3	2338	12	Q14183	N.clavipes draglin
20	56	1.3	16442	18	X83006	Partial mouse WRN
21	55.8	1.3	2004	18	T85356	Nephila clavipes s
22	54.6	1.3	32207	20	V73805	KSHV LUR DNA (nucl
23	54.6	1.3	137507	19	V19941	KSHV long unique c
24	54.4	1.3	9551	20	Z22301	cDNA encoding a hu
25	54	1.3	543	13	Q23092	Antigen tc-7a gene
26	53.6	1.3	1995	12	Q14184	N.clavipes draglin
27	53.6	1.3	1995	19	V23250	Nephila clavipes s
28	53.6	1.3	1995	21	Z38196	N. clavipes spider
29	53.4	1.2	4055	20	Z40484	Human ZC2 DNA. Ho
30	53.4	1.2	4180	21	A10669	Human protein kina
31	53	1.2	1137	14	Q33061	Plasmodium vivax c
32	52.6	1.2	10732	21	A10594	Gene encoding a su
33	52.4	1.2	114955	20	X53491	Human adenosine A1
34	52.2	1.2	1847	19	V04694	Cell cycle protein
35	51.4	1.2	3198	20	X02974	Human IL-1ra BAC c
36	51.2	1.2	2744	16	Q98470	MisP1-containing p
37	51.2	1.2	51259	18	X83007	Partial mouse WRN
38	50.2	1.2	1686	16	Q87587	DNA encoding Leuco
39	49	1.1	2040	18	T62137	Leishmania brazili
40	49	1.1	2040	19	V47557	Leishmania antigen
41	48.8	1.1	1218	21	A02488	Human colon cancer
42	48.8	1.1	2313	20	Z07197	Human lung tumour
43	48.6	1.1	1432	19	V47585	Leishmania antigen
44	47.8	1.1	3337	17	T34620	P. vivax ESP-1 blo
45	47.8	1.1	3337	20	X15174	DNA encoding a sec

ALIGNMENTS

RESULT 1
T76739
ID T76739 standard; cDNA; 2647 BP.
XX
AC T76739;
XX
DT 08-OCT-1997 (first entry)
XX
DE cDNA encoding active type R-Ras protein binding protein p98.
XX
KW Active-type-R-Ras protein; p98; tumour; disease; cancer;
KW cellular Ca ion; phospholipid; tumour forming promoter;
KW tumour forming inhibitor; ss.
XX
OS Bos taurus.
XX
FH Key
FT CDS
FT
FT
FT
XX
PN JP09135688-A.
XX
PD 27-MAY-1997.
XX
PF 16-NOV-1995; 95JP-0298720.
XX
PR 16-NOV-1995; 95JP-0298720.
XX
PA (KIRI) KIRIN BREWERY KK.
XX
DR WPI; 1997-335993/31.
DR P-PSDB; W24227.
XX
PT Active R-Ras protein-binding protein p98 - used in the treatment of

PT cancer, and diseases related to cellular calcium ion concentrations
XX
PS
XX Claim 9; Page 9-12; 18pp; Japanese.
CC This sequence encodes a protein capable of binding to active-type-R-Ras
CC protein. The protein has a molecular weight of 98 kD by SDS-PAGE, and
CC is designated p98. The protein coding sequence can be inserted into a
CC suitable vector to treat certain tumours or diseases (various cancers)
CC in which active R-Ras protein is involved. Because active R-Ras protein
CC is involved in formation and inhibition of tumours and related to
CC diseases in which cellular Ca ions or phospholipids are involved, the
CC screening method can be used for screening e.g. a tumour forming promoter
CC or tumour forming inhibitor.
XX
SQ Sequence 2647 BP; 633 A; 764 C; 769 G; 481 T; 0 other;

Query Match	2.2%;	Score 95.8;	DB 18;	Length 2647;
Best Local Similarity	50.9%;	Pred. No. 1.7e-12;		
Matches 310; Conservative	0;	Mismatches 287;	Indels 12;	Gaps 3;

1278	QY	caagggaagaggaggtcgctagtcgactgggttcacatcctgcacaaagcacagggaagggc	1337
996	Db	cagggaagacgaggagcgccatcccgctggctgctcttctgcattacggccgggt	1055
1338	QY	caaggacttctttcagacatggccatgtcagaggtagaccgggttcacggagcggaaca	1397
1056	Db	ggtgcccttcacagcgccatcgccagcgccgaggtccgaggacccaggacc--ccaa	1112
1398	QY	cctcatattccgcgagaacacgctcgccactaaagccatagaagagtatatgagactgat	1457
1113	Db	tacgatcttcgagggaactcgctgacatccaagtgcacatgcagtgcagcagatgaagctggc	1172
1458	QY	tggccagaaaatacctcaaggatgccattggggagttcatccgggctctgtatgaatctga	1517
1173	Db	aggcatgcagtagcctacacgtcacccctgaacccaccacatagaggagatttgcacagagtca	1232
1518	QY	ggagaactgtgaagtagaccccccatcaagtgcacagcgtcc--agtcctggcagagcacca	1574
1233	Db	taagtctgtgaaattgaccccgtagggctgaaggacggtcagagcctagagagcaacat	1292
1575	QY	ggccaactgcggatgtgctgtgagttggccctgtgcaaggtggtcaactcccattgctg	1634
1293	Db	ggagaacctgcggcagttgtggaccgcgtgttcagcgtcatcaccaaatcgggggtgag	1352
1635	QY	gttcccgaggagctgaaggaggtgtttgcacatcatggcggctgcgtgtgcagagcgggg	1694
1353	Db	ctgccccaccgtcatgtgtacatcttcttctccctcggggagcgccgcacagcgctt	1412
1695	QY	ccgggaggacattgctg-----acaggctgatcagcgccctgcgtcttcttctgcgttctct	1748
1413	Db	ccaggatgaacctggacgtgcggtacacagctgtgagcagcttcatcttctgcggttctt	1472
1749	QY	ctgcccgccatcatgtcgccagctctgtttggactgatgcaggagtacccagatgagca	1808
1473	Db	cgctccagccatcctgtcccccaacctcttccagctcacgccccaccacacggacccaca	1532
1809	QY	gacctcacgaacctcacccctcatcgcccaaggttatccagaacctggccaaacttttccaa	1868
1533	Db	gacgtcgaggacgctgacacttgttttcgaagaccattcagactctggtggcagcctgtccaa	1592
1869	QY	gtttacctc	1877
1593	Db	gtccaa	1601

RESULT	2
Z17263	
ID	Z17263 standard; cDNA; 1337 BP.
XX	
AC	Z17263;
XX	
DT	12-OCT-1999 (first entry)

Human gene expression product cDNA sequence SEQ ID NO:4735.

Human; gene; gene expression product; diagnosis; therapy; probe; detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

Homo sapiens.

WO9938972-A2.

05-AUG-1999.

28-JAN-1999; 99WO-US01619.

03-APR-1998; 98US-0080666.

28-JAN-1998; 98US-0072910.

24-FEB-1998; 98US-0075954.

31-MAR-1998; 98US-0080114.

03-APR-1998; 98US-0080515.

(CHIR) CHIRON CORP.
(HYSE-) HYSEQ INC.

Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
Stache-Crain B, Sudduth-Klinger J, Williams LT;

WPI; 1999-494092/41.

Novel human genes and their expression products which are differentially expressed in different cell types

Claim 1; Page 2250-2251; 2479pp; English.

The present invention describes a library of human polynucleotides comprising the sequences given in Z12532 to Z17779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5248 polynucleotide sequences given in Z12532 to Z17779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purpose, e.g. detection of expression levels, mapping, tissue typing or profiling, forensics, genetic analysis and detection of polymorphisms. Polypeptides encoded by the polynucleotides can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the diagnosis, prognosis and management of colorectal cancer, breast cancer, and lung cancer. The polynucleotides can also be used to screen for peptide analogues and antagonists.

Sequence 1337 BP; 42A; 577 C; 27 G; 22 T; 669 other;

Query Match	1.5%;	Score 63.6;	DB 20;	Length 1337;
Best Local Similarity	33.0%;	Pred. No. 2.7e-05;		
Matches 147; Conservative	0;	Mismatches 298;	Indels 0;	Gaps 0;

[illegible]

Best 11.1%

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2001, 01:35:57 ; Search time 6516.34 Seconds
(without alignments)
4593.988 Million cell updates/sec

Title: US-09-294-298-1
Perfect score: 4272
Sequence: 1 atgtctatgcccccttcag.....tgggggtgctatccccatcct 4272

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues
Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_estl1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *
10: gb_est10: *
11: gb_est11: *
12: gb_est12: *
13: gb_est13: *
14: gb_est14: *
15: gb_est15: *
16: gb_est16: *
17: gb_est17: *
18: gb_est18: *
19: gb_est19: *
20: gb_est20: *
21: gb_est21: *
22: gb_est22: *
23: gb_est23: *
24: gb_est24: *
25: gb_est25: *
26: gb_est26: *
27: gb_est27: *
28: gb_est28: *
29: gb_est29: *
30: gb_est30: *
31: gb_est31: *
32: gb_est32: *
33: gb_est33: *
34: gb_est34: *
35: gb_est35: *
36: gb_est36: *
37: gb_est37: *
38: gb_est38: *
39: gb_est39: *
40: gb_est40: *
41: gb_estba: *
42: em_estfun: *
43: em_esthum1: *

44: em_esthum2: *
45: em_esthum3: *
46: em_esthum4: *
47: em_esthum5: *
48: em_esthum6: *
49: em_esthum7: *
50: em_esthum8: *
51: em_esthum9: *
52: em_esthum10: *
53: em_esthum11: *
54: em_esthum12: *
55: em_esthum13: *
56: em_esthum14: *
57: em_esthum15: *
58: em_esthum16: *
59: em_esthum17: *
60: em_esthum18: *
61: em_esthum19: *
62: em_esthum20: *
63: em_estini1: *
64: em_estini2: *
65: em_estini3: *
66: em_estini4: *
67: em_estov1: *
68: em_estov2: *
69: em_estpl1: *
70: em_estpl2: *
71: em_estpl3: *
72: em_estpl4: *
73: em_estpl5: *
74: em_estro1: *
75: em_estro2: *
76: em_estro3: *
77: em_estro4: *
78: em_estro5: *
79: em_estro6: *
80: em_estro7: *
81: em_estro8: *
82: em_estro9: *
83: em_estro10: *
84: em_estro11: *
85: em_estro12: *
86: em_estro13: *
87: gb_est41: *
88: gb_est42: *
89: gb_est43: *
90: gb_est44: *
91: gb_est45: *
92: gb_est46: *
93: gb_est47: *
94: gb_est48: *
95: gb_est49: *
96: gb_est50: *
97: gb_est51: *
98: gb_est52: *
99: gb_est53: *
100: gb_est54: *
101: gb_est55: *
102: gb_est56: *
103: gb_est57: *
104: gb_est67: *
105: gb_est68: *
106: gb_est69: *
107: gb_est70: *
108: gb_est71: *
109: gb_est72: *
110: gb_est73: *
111: gb_est74: *
112: em_esthum21: *
113: em_esthum22: *
114: em_esthum23: *
115: em_estom1: *
116: em_estom2: *

117: em_estpl6:*
118: em_estpl7:*
119: em_estpl8:*
120: em_estpl9:*
121: em_estpl10:*
122: em_estpl11:*
123: em_estpl12:*
124: em_estpl13:*
125: em_estpl14:*
126: gb_est58:*
127: gb_est59:*
128: gb_est60:*
129: gb_est61:*
130: gb_est62:*
131: gb_est63:*
132: gb_est64:*
133: gb_est65:*
134: gb_est66:*
135: gb_est67:*
136: gb_est68:*
137: gb_est69:*
138: gb_est70:*
139: gb_est71:*
140: gb_est72:*
141: gb_est73:*
142: gb_est74:*
143: gb_est75:*
144: gb_est76:*
145: gb_est77:*
146: gb_est78:*
147: gb_est79:*
148: gb_est80:*
149: gb_est81:*
150: gb_est82:*
151: gb_est83:*
152: gb_est84:*
153: gb_est85:*
154: gb_est86:*
155: gb_est87:*
156: gb_est88:*
157: gb_est89:*
158: gb_est90:*
159: gb_est91:*
160: gb_est92:*
161: gb_est93:*
162: gb_est94:*
163: gb_est95:*
164: gb_est96:*
165: gb_est97:*
166: gb_est98:*
167: gb_est99:*
168: gb_est100:*
169: gb_est101:*
170: gb_est102:*
171: gb_est103:*
172: gb_est104:*
173: gb_est105:*
174: gb_est106:*
175: gb_est107:*
176: gb_est108:*
177: gb_est109:*
178: gb_est110:*
179: gb_est111:*
180: gb_est112:*
181: gb_est113:*
182: gb_est114:*
183: gb_est115:*
184: gb_est116:*
185: gb_est117:*
186: gb_est118:*
187: gb_est119:*
188: gb_est120:*
189: gb_est121:*

190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	490.4	11.5	571	37	AV590698	AV590698 AV590698
2	383.8	9.0	474	181	AZ399131	AZ399131 LM0164P14
3	294.6	6.9	593	95	AW779747	AW779747 hn85f10.x
4	250.4	5.9	463	108	BE463433	BE463433 hw23d06.x
5	244.6	5.7	427	11	AA709010	AA709010 zf94h11.s
6	233.6	5.5	597	90	AW421227	AW421227 fj94f10.x
7	228.8	5.4	682	173	AQ993964	AQ993964 RPCI-23-2
8	222	5.2	566	22	AI606711	AI606711 ml58d01.y
9	221.8	5.2	413	10	AA691827	AA691827 vt05d01.r
10	214	5.0	707	28	AJ396547	AJ396547 AJ396547
11	211.8	5.0	555	27	AI958829	AI958829 fd22901.y
12	201	4.7	628	108	BE490923	BE490923 db38g08.x
13	197.2	4.6	516	111	BE695645	BE695645 MRI-BT080
14	195.8	4.6	347	87	AW204612	AW204612 UI-H-BI1-
15	195.4	4.6	596	136	BE848541	BE848541 uw39f05.y
16	190	4.4	481	135	BE773015	BE773015 RC1-FT013
17	185	4.3	542	93	AW656799	AW656799 109245 MA
18	182.2	4.3	296	23	AI650331	AI650331 wa18f01.x
19	177.6	4.2	345	39	AW076911	AW076911 fj03d08.y
20	173	4.0	511	22	AI577567	AI577567 UI-R-Y0-V
21	167.2	3.9	388	183	B68206	B68206 CIT978SK-A-
22	165.6	3.9	474	87	AW205989	AW205989 UI-H-BI1-
23	161.2	3.8	460	22	AI609604	AI609604 tw91e09.x
24	157.8	3.7	1020	191	CNS02VGV	AL215753 Tetraodon
25	157	3.7	449	9	AA554056	AA554056 n101g07.s
26	154.6	3.6	440	2	AA107246	AA107246 ml58d01.r
27	149.4	3.5	651	19	AI327335	AI327335 mp74c11.x
28	148.6	3.5	877	192	CNS04DL0	AL285885 Tetraodon
29	146.8	3.4	443	40	AW136165	AW136165 UI-H-BI1-
30	146.8	3.4	464	95	AW826497	AW826497 fk62h05.x
31	146	3.4	452	9	AA554055	AA554055 n101g06.s
32	145.4	3.4	407	133	BE015978	BE015978 fk62h05.y
33	139.8	3.3	382	22	AI562171	AI562171 vw73f04.x
34	138.6	3.2	317	91	AW527364	AW527364 UI-R-B01-
35	138.6	3.2	457	2	AA118397	AA118397 mp74c11.r
36	136.4	3.2	268	92	AW593427	AW593427 hgl5e10.x
37	136.6	3.2	379	9	AA575948	AA575948 nm56e05.s
38	135.6	3.2	554	111	BE723198	BE723198 192621 MA
39	133.4	3.1	336	26	AI874961	AI874961 ul27f03.x
40	131.8	3.1	694	28	AJ399084	AJ399084 AJ399084
41	130.4	3.1	335	134	BE085114	BE085114 CM2-BT066
42	130	3.0	1006	190	CNS00HWF	AL073826 Drosophil
43	128.2	3.0	631	191	CNS01XQ5	AL172022 Tetraodon
44	126.8	3.0	975	191	CNS03ANH	AL235430 Tetraodon
45	124.2	2.9	397	111	BE695651	BE695651 MRI-BT080

ALIGNMENTS

RESULT 1
AV590698
LOCUS AV590698 571 bp mRNA
DEFINITION AV590698 Bos taurus brain fetus Bos taurus cDNA clone E1BR014B04
5', mRNA sequence.
ACCESSION AV590698
VERSION AV590698.1 GI:9701691
KEYWORDS EST.
SOURCE COW.

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 571)
AUTHORS Sugimoto, Y., Hirotsune, S., Takasuga, A., Itoh, R., Jitohzono, A. and Suzuki, H.

TITLE bovine cDNA sequencing
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@coea.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES Location/Qualifiers

1..571
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="E1BR014B04"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A was deleted from a NotI site"

BASE COUNT 124 a 161 c 169 g 116 t 1 others
ORIGIN

Query Match 11.5%; Score 490.4; DB 37; Length 571;
Best Local Similarity 93.2%; Pred. No. 8.7e-104;
Matches 523; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 1283 gcaaggaggaggtcgtagtgactggttcacatctcgtcaagagcagcagggccaagg 1342
DB 11 GCAAGGAGGAGGTCGCCAGTGCTCTGTGTTACATCTCTGAGAGTACAGGCAAGGCAAGG 70
QY 1343 acttctcttcagacatggtccatgtagaggttagaccgttcagcggtagcgggaacacctca 1402
DB 71 ACTTCTTTTCAGACATGGCCATGTCAGAGGTGGACCGGTTCATGGAACGAGAACACCTCA 130
QY 1403 tattccgcgagaaacagctcgccactaagccatagagagtagatgagactgattggcc 1462
DB 131 TATTCGCGGAGAACACGCTCGCCACTAAAGCCATAGAGAGTACATGAGACTGATTGGTC 190
QY 1463 agaaataacctcaaggatgcccattggggaggttcattccgggctctgtatgaatctgaggaga 1522
DB 191 AGAAATACCTCAAGGATGCCATCGGGGAATTATCATCCGTGCTCTGTATGAATCTGAGGAGA 250
QY 1523 actgtgaagttagaccccatcaagtgcacagcgtccagtcgtgagagcagcagcagcagcagc 1582
DB 251 ACTGTGAGGTGGACCCCATCAAGTGCACGGCATCCAGTTGGCTGAGCAGCAGGCAACC 310
QY 1583 tgcggatgtgctgtgagttggccctgtgcaaggttgtaactcccatcctggtgttcccca 1642
DB 311 TGCGGATGTGCTGTGAGTGGCCCTGTGCAAGGTGGTCACTCCCATTCGCTGTTCCTCCA 370
QY 1643 gggagctgaaggaggtgttgcatcatggtcggtcggtgagcagcagcagcagcagcagcagc 1702
DB 371 GGGAGCTGAAGGAGGTGTTCGCTCTGCGGACTGCGGTGTGTCAGAGCGGGCGCGAGAGG 430
QY 1703 acattgtgacaggtgatcagcgcctcgctcttccgtgcttccctcctcctcctcctcctc 1762
DB 431 ACATTGCGGACCGGCTGATCAGCGCCTCGCTCTTCTCTGCGCTTCTCTGCGGCGGCAATTA 490
QY 1763 tgtcgccagctgtgtt-tggactgatcaggagtagaccagatgagcagcagcagcagcagc 1821
DB 491 TGTCGCCACGCTNTTCGGGGCTTCATGAGGAGGTACCCAGATGAGCAAACTTCGAGAAC 550

QY 1822 ctcaccctcatcgcccaagggtt 1842
DB 551 CTCACGCTCATCGCCCAAGGTT 571

RESULT 2
AZ399131/C
LOCUS 474 bp DNA GSS 03-OCT-2000
DEFINITION LM0164P14R Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCLM0164P14 R, DNA sequence.
ACCESSION AZ399131
VERSION AZ399131.1 GI:10514203
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 474)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0164 row: P column: 14
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 474.

FEATURES Location/Qualifiers
1..474
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0164P14"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42hv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 92 a 151 c 131 g 100 t
ORIGIN

Query Match 9.0%; Score 383.8; DB 181; Length 474;
Best Local Similarity 93.7%; Pred. No. 5.9e-79;
Matches 400; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

